AD-A202 238

:UMENTATION PAGE

DTIC FILE COPY

Unclassified	**1b. RESTRICTIVE MARKINGS** --	
2a. SECURITY CLASSIFICATION AUTHORITY --	**3. DISTRIBUTION/AVAILABILITY OF REPORT** Approved for public release; distribution unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) --	**5. MONITORING ORGANIZATION REPORT NUMBER(S)** ARI Research Note 88-78	

6a NAME OF PERFORMING ORGANIZATION Florida Atlantic University	**6b. OFFICE SYMBOL** (If applicable) --	**7a. NAME OF MONITORING ORGANIZATION** U.S. Army Research Institute for the Behavioral and Social Sciences
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Boca Raton, Florida 33431		**7b. ADDRESS (City, State, and ZIP Code)** 5001 Eisenhower Avenue Alexandria, VA 22333-5600
8a. NAME OF FUNDING/SPONSORING ORGANIZATION --	**8b. OFFICE SYMBOL** (If applicable) --	**9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER** MDA903-82-C-0317

8c. ADDRESS (City, State, and ZIP Code) --

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
6.11.02	B74F	n/a	n/a

11. TITLE (Include Security Classification)

Frequency Monitoring: A Methodology for Assessing the Organization of Information

12. PERSONAL AUTHOR(S)
Howard S. Hock and Lynn Hasher

13a. TYPE OF REPORT Interim Report	**13b. TIME COVERED** FROM 5/85 TO 9/86	**14. DATE OF REPORT (Year, Month, Day)** 1988, August	**15. PAGE COUNT** 189

16. SUPPLEMENTARY NOTATION

Michael Drillings, contracting officer's representative

17.	COSATI CODES		**18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)**
FIELD	GROUP	SUB-GROUP	Encoding Strategies Spatial Relations Cognitive Psychology Judgement Information Theory Memory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The general purpose of the research reported in this research note was to evaluate the viability of a technique for assessing the encoding of componential and higher-order memory units of visual stimuli. The first section of the note provides an extended summary of results for encoding strategies, spatial relations research, and frequency judgement theory. The second section describes experiments which specifically use the frequency-judgement procedure, and section three describes experiments conceptually related to the frequency-judgement work.

DTIC
ELECTE
DEC 30 1988

88 12 29 083

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	**21. ABSTRACT SECURITY CLASSIFICATION** Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Michael Kaplan	**22b. TELEPHONE (Include Area Code)** 202/274-8722	**22c. OFFICE SYMBOL** BRO

DD Form 1473, JUN 86   Previous editions are obsolete.   SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

CONTENTS

Frequency Monitoring: A Methodology for Assessing the Organization of Information

Howard S. Hock and Lynn Hasher
Florida State University

for

Contracting Officer's Representative
Michael Drillings

ARI Scientific Coordination Office, London
Milton S. Katz, Chief

Basic Research Laboratory
Michael Kaplan, Director

U. S. Army
Research Institute for the Behavioral and Social Sciences

August 1988

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Research accomplished under contract
for the Department of the Army

Florida Atlantic University

Technical review by

Dan Ragland

INTRODUCTION

The general purpose of the research supported by this grant was to evaluate the viability of a technique for assessing the encoding of componential and higher-order memory units of visual stimuli. The technique involves presented a sequence of stimuli such that certain visual information is presented with different frequencies. For example, the frequency with which different words occur in a list of words might vary, as might the frequency of occurrence of different letters constituting the words. The presentation of the stimulus sequence is followed by a memory test in which subjects judge the frequency with which specified events occurred in the sequence. In applying this procedure to a variety of stimulus materials and tasks, we have been able to assess the extent to which subjects' judgments of frequency of occurrence are influenced by both component-level and higher-order memory units. In addition to having many practical advantages, the frequency-judgment technique appears to be a valuable tool for assessing certain kinds of learning and answering theoretical questions relating to the encoding of memory units.

The first section of the report provides an extended summary of the results obtained in the grant research. More detailed accounts of experimental procedures and results follow: Section 2 describes experiments that specifically use the frequency-judgment procedure, and Section 3 describes experiments that are conceptually related to the frequency-judgment work (including some research initiated in conjunction with an earlier ARI grant to the principal investigator).

Section 1
SUMMARY OF FINDINGS

A. Practical Advantages of the Frequency-Judgment Technique

a. _Incidental Learning._ Accurate frequency judgments can be obtained under purely incidental learning conditions. Although researchers have been concerned with the theoretical implications of small differences in the accuracy of frequency judgments under intentional and incidental conditions (Greene, 1984; Moshe-Benjamin & Jonides, 1986), these differences are probably of little practical importance. The opportunity for incidental learning makes the technique particularly useful for individuals who react negatively when placed in situations in which learning something is explicitly required of them. Along the same lines, the use of incidental tasks with measurable outcomes allows one to be certain that the individual is paying attention to the material that is presented to them. For example, Hock, Malcus, and Hasher (1986; Section 2A) have used a lexical decision task to obtain incidental learning of letter-level and string-level frequency information. The lexical decision task task does not have an explicit memory requirement; subjects are simply required to judge whether or not a string of letters constitutes an English word. Subjects' attentiveness in this task can be ascertained by examining the speed and accuracy of their responses and determining whether they are within a normal range.

b. _Performance Is Not Affected by Individual Differences._ Previous research, summarized by Hasher and Zacks (1979), has demonstrated that frequency-judgment techniques are relatively unaffected by the individual difference variables that heavily influence other measures of memory. Thus, it is relatively unaffected by such factors as age, motivation, intelligence, and anxiety. This is not to say that such factors never influence the accuracy of frequency judgments. The point that is clear is that the effects of such factors are relatively small compared with thier impact on other memory techniques. From a practical point of view, this means that there is less need to tailor the learning context to take account of individual differences.

c. _It is an Ideal Device for Studying Memory for Componential Units._ As pointed out by Hock, Malcus, and Hasher (1986; Section 2A), the frequency-judgment procedure is particularly useful for assessing the encoding of element-level memory units when the vocabulary of elementary units is limited (e.g., there are only 26 alphabet letters). Thus, when frequency judgments are used to assess memory for the constituent letters in strings, the full set of alphabet letters can be presented during the acquisition phase of the experiment. In contrast, recognition and recall procedures would require "holding back" a significant proportion of the letters from the acquisition phase because they must be used to detect false recognition responses or recall intrusions during the subsequent memory test (imagine if all the letters were presented during acquisition; subjects could accurately "recall" them simply by reciting the entire alphabet).

d. _It Eliminates the Problem of Production Deficits._ In the absence of well-developed strategies for retrieving information from

memory, tasks requiring subjects to explicitly state what they remember often underestimate what they have learned. Production deficits can also be obtained when individuals lack confidence in what they know; they will tend to withhold responses when they must produce a description of what they've learned. The frequency-judgment technique minimizes the need for production. Subjects will demonstrate knowlege of frequency of occurrence information even under conditions in which they claim they recall nothing, and for that matter, under conditions in which they claim they do not remember the frequency of the events that they are asked to judge. The frequency-judgment procedure is like recognition testing in that the item being tested is provided to the learner; they need only judge how often it occurred (in recognition testing the learner need only indicate whether or not the item was seen before).

e. It Eliminates the Problem of Distractor Similarity. Although it is like recognition testing in that it eliminates the need for production, the frequency-judgment technique has an important advantage compared with recognition techniques. This is because recognition testing could not proceed simply by presenting previously seen items and asking subjects whether or not they look familair. Someone with no memory of the items may simply decide to say "Yes" to each of them, giving the erroneous appearance of perfect recognition. For this reason, recognition testing procedures do not test for whether an item looks familiar. Rather, they test subjects' ability to discriminate between previously seen and new items (the later are referred to as distractors). The problem is that recognition performance is critically dependent on the similarity of the previously learned items to the distractor items included in the list. Since recognition accuracy will decrease as the similarity of the distractors to the previously learned items increases, recognition performance is not a true estimator how well someone has remembered previously learned information; different estimates of memory accuracy are obtained for different sets of distractors. This problem does not arise in the frequency-judgment paradigm since it does not require the presentation of distractors. Subjects are tested only with the items that they are presumed to have learned during acquisition.

B. Subjects Can Judge the Frequency of Occurrence for Component Elements of Higher-Order Perceptual Units

The experiments we've performed have shown that subjects can judge the frequency with which letters appear in a sequence of words, the frequency with which words appear in a sequence of sentences, the frequency with which locations within a frame are occupied by the elements composing a sequence of patterns, and the frequency with which various spatial relations appear in a sequence of scenes. During the first, or input phase of our experiments, subjects were presented a series of stimuli under various instructional conditions. This was always followed by a second phase, in which they were asked to judge the frequency with which various events occurred during the first phase. Most of the time, the instructions accompanying the first phase were incidental to the subsequent frequency-judgment task. Thus, after seeing a sequence of words and nonwords, subjects were surprised when they were asked to estimate how

often various letters occurred over the sequence of letter strings. After seeing a sequence of patterns, the constituent elements of which were circles inside a square frame, they were similarly surprised when they were asked to estimate how often a circle appeared at various locations inside the frame. Subjects performance in these tasks is quite remarkable in the context of their virtually universal insistence that they could not perform the judgment task. They had to be assured that people always do much better than they expect at this sort of task, and they were encouraged to try hard and give their best guess. Despite being certain that they could not perform the task, when coaxed to do so subjects successfully judged frequency of occurrence for the components of words, patterns and scenes.

On an individual level, performance was not always very good. Estimation accuracy, which was measured by computing correlation coefficients between actual and estimated frequency, was very variable. Subjects sometimes had correlations that were quite high (in the 0.90s), but they sometimes had strongly negative correlations (e.g., -0.61). The average of individual subject correlations was as low as 0.17 (in one of the conditions involving the estimation of element-location frequency for patterns), but sometimes were quite good (as high as 0.56 in one of the conditions involving the estimation of letters occurring in strings). Since the correlation between actual and estimated frequency was based on as few as eight (and at most 16) data points, the results obtained for individual subjects were highly sensitive to many factors that could not be adequately controlled. When frequency estimates of componential information were averaged over all the subjects participating in a condition, correlations between mean-estimated and actual frequency were substantially larger than the means of individual correlations.

It was with respect to these correlations between mean-estimated and actual frequency that the estimation technique demonstrated its greatest practical viability. These correlations ranged from the 0.60s to the 0.90s, depending on the condition. From a practical point of view, the pooling of estimation data from groups with 16 to 24 individuals appears to be sufficient to obtain high accuracy in the estimation of component-level information. One direction of future research might be to examine the social-dynamic factors that would influence pooled-decision making in frequency estimation tasks.

C. Estimates of Componential Frequency Are Based on Component-Level Memory Units

One of the first issues we addressed in our experiments was what sort of memory units subjects use as the basis for their judgments of component-level frequency information. One possibility was that subjects base their estimates on the retrieval of component-level memory units that were abstracted over the full sequence of items (letter strings or patterns) that were presented. The second possibility, which is related to the "availability heuristic" proposed by Tversky and Kahneman (1973), is that subjects derive their estimates of component information by retrieving global-level memory units that include the component being judged. The more such global units that are available (i.e., retrievable), the higher the frequency estimate will be for any component. This is a viable decision rule

because a high frequency component occurs in a large number of
different global units, or else it appears in global units that occur
with high frequency. In either case, the probability of retrieving
global-level units comprising the component-level unit is enhanced.

Hock, Malcus, and Hasher (1986; Section 2A) used several
techniques to demonstrate that frequency estimation for letter-level
information can be based almost exclusively on component-level codes.
One of the primary techniques involves having subjects recall as many
strings as possible from those that they had just seen (this was done
immediately after they judged frequency of occurrence for the
component letters of the strings). The assumption we made here was
that the likelihood of a string being recalled was directly related to
the likelihood of its memory representation being activated while
subjects were making judgments of letter frequency. Based on this
assumption, we counted the number of times each letter appeared in the
strings subjects correctly recalled (i.e., the recall-frequency for
each letter), and inferred that this represented what the subject's
estimate for each letter would be if estimates were based on
"availability," indexed here by recall of strings that included the
letter. We then computed a partial correlation coefficient in which
the influence of recall-frequency was "held constant." The
correlations between estimated and actual frequency remained
significantly positive, demonstrating that subjects had information
available on the frequency of components that was at least partially
independent of their knowledge of the global units (i.e., strings).

Similar results were also obtained by Hock and LaLomia (Section
2B) in an experiment in which subjects were presented a sequence of
sentences and asked to judge the frequency with which various words
appeared over the full set of sentences. Correlations between
estimated and actual frequency remained significantly positive after
the frequency with which the words appeared in correctly recalled
sentences (or even incorrectly recalled sentences) was "partialled
out." Additional experiments demonstrated the same for subjects'
knowledge of the frequency with which locations within a frame were
occupied by the elements constituting a series of patterns (Hock,
Smith, Escoffery & Bates, 1985, Section 2C). That is, subjects'
knowledge of location frequency was at least partially independent of
their knoledge of the global units (i.e., the patterns).

D. Concerning the Issue of Automaticity

Although it was not a specific objective of the grant to study
whether frequency of occurrence information is encoded automatically,
the issue has become sufficiently controversial (Moshe-Benjamin &
Jonides, 1986; Greene, 1984; Fisk, 1986) to impact on all work
concerned with frequency judgments. The argument that frequency
information is encoded automatically was developed in two papers by
Hasher and Zacks (1979; 1984). Among several criteria for assessing
automaticity, the one that has received the most attention has
concerned the influence of the subject's "cover" task during the
initial presentation of the items. According to the strictest
interpretation of automaticity, the accuracy of frequency judgments
should not depend on the subject's task, whether the contrast is
between intentional or incidental cover tasks, or between two types of

incidental task. Experiments by Moshe-Benjamin and Jonides (1986) and Greene (1984) have challenged this interpretation; they've obtained differences in the accuracy of frequency judgment under different task conditions. Another approach to the issue of automaticity emphasizes the effortless nature of frequency encoding. Although general instructions to remember the items in a list can lead to more accurate frequency judgments than instructions with no memory requirement (Hock & Cavedo, Section 2D), the fact that subjects are successful at encoding frequency under purely incidental conditions (Hasher, Zacks, Rose, & Sanft, 1987) indicates that frequency information is encoded without any effort to do so. However, the implication of this view, that frequency encoding is inevitable, has been challenged by research showing that subjects given extensive practice can make semantic decisions about words, but retain no memory of the frequency of occurrence of these words (Fisk & Schneider, 1984).

A way of clarifying these conflicting claims has been suggested by Zacks, Hasher, and Hock (1986; Section 2E). Referring back to Hasher and Zacks' (1979) assertion that frequency of occurrence information is encoded automatically only for stimuli that receive attention, Zacks, Hasher, and Hock (1986) elaborated on the attentional aspect of automaticity in conjunction with "late selection" views of attention, for example, Duncan's (1980). Duncan has argued that stimuli are fully analyzed preattentively (i.e., without the use of attentional resources), including the extraction of their form and meaning. The limited capacity attentional system determines which products of preattentive processing will be attended to and thereby brought into consciousness. It follows from Duncan's model that evidence a stimulus is processed, even semantically, does not constitute evidence that the stimulus has been attended to. Fisk and Schneider (1984) trained subjects to the point where conscious attention was not required for subjects to make semantic decisions. With no attention, there was no coding of frequency information.

In a similar vein, differences in judgment accuracy following different cover tasks might also be attributable to differences in attention to the items as they are processed. To demonstrate this point, Hock and LaLomia (Section 2B) showed that frequency judgments for nouns that appeared in a sequence of sentences were more accurate when subjects' cover task was to determine whether or not the sentences were meaningful compared with when their cover task was to determine whether the sentence was written in present or past tense. We concluded that relatively inaccurate judgment of word frequency for the nouns in the tense-judgment condition was the result of subjects directing their attention to the verbs in each sentence [there was some processing of the semantic content of the sentences in the tense-judgment condition; judgments of verb tense were faster for verbs appearing in meaningful sentences than for verbs appearing in meaningless sentences].

But why is attention important? In an experiment with sequences of words rather than sequences of sentences, Malcus, Hock, and Cavedo (1985; Section 2F) had one group of subjects participate in a lexical decision task, the other in a letter search task. For word lists in which each word appeared only once, we found a vast advantage in recall accuracy for the lexical decision condition compared with the letter search condition. However, when each word appeared more than

once (the frequency of each word varied), there was no difference in recall accuracy or frequency-judgment accuracy for the two conditions. These results suggest that the presence of repetitions may affect the size of the perceptual unit to which subjects attend. When subjects are looking for certain target letters (a U or I) in the letter search task, they are likely to focus their attention on individual letters. Hence, there is relatively little encoding of word-level memory units. However, when the same word reappears many times, the likelihood increases that subjects will recognize a word (e.g., "PINE") as a word they previously made a "Yes" response to (because there was an I in it), and repeat that response. Thus, the presence of repetitions can shift subjects' attention from letter-level units to word-level units.

But why did subjects shift their attention to the word-level when the task ostensibly required letter-level processing? The apparent reason was that efficient processing could be achieved by searching backwards in memory and reactivating previously stored words (e.g., "PINE") and their associated responses (the "Yes" response for "I"). It is our current view that this backward search could not be done without attention, and it is the resultant reactivation of word-level codes corresponding to previously seen words that facilitates frequency judgments for the words (probably by the creation of more traces). From this point of view, differences in the accuracy of frequency judgments for different cover tasks are the result of differences in the extent to which there are attention-dependent backwards searches through previously presented items in a list. Thus, the encoding of information that is the basis of accurate frequency judgments occurs without intention or effort (and is automatic in that sense), but effortful backward searches (for the purpose of improving performance in the cover task) can improve the accuracy of the frequency judgments.

A series of experiments to test the backward-search hypothesis could be designed in the context of van Dijk 'and Kintsch's (1983) model of text comprehension. One of the basic principles in the model is that as we read text, each item (sentence) is comprehended in relation to previously encoded information in the text. Some of this previously encoded information is kept in an active buffer; this is usually information of central importance to the text as well as recently encoded information. The remaining information in the text is stored in long-term memory. When subjects read a sentence, they first seek "argument overlap" with information that is active in the buffer. If overlap is obtained for items in the buffer, there would be no additional activation of memory traces because the overlapping information is already active in memory. If overlap is not obtained with information in the buffer, it must result from backward search through long-term memory. The information in long-term memory will be then be reactivated, producing additional memory traces that would facilitate subsequent frequency of occurrence judgments.

The experiments would therefore test the following hypotheses: 1) the introduction of sentences in a text that requires reactivating more instances of a previously presented word in long-term memory should facilitate subsequent frequency judgments of that word, and 2) the introduction of sentences that can be understood in conjunction with information in the buffer should decrease the accuracy of frequency judgments. The latter result would provide an explanation

of the "spacing effect" on frequency judgments (Hintzman & Block, 1971). That is, frequency judgments would be better when repeated items are spaced far apart because backward search through long-term memory would reactivate previously formed traces of the item; when the items are closely spaced, previous instances would be likely to remain active in the buffer, so there would be no backward search to reactivate the previously formed traces of the item (and hence, no facilitation of frequency judgments.

E. Evidence for Letter-Level Recognition Units for Words and Acronyms

When subjects are required to indicate whether or not a target letter is included in a briefly presented string of letters, they are more accurate when the string is a familiar word than when it is a nonword (e.g., Reicher, 1969). This has result is known as the word superiority effect (WSE). Additional experiments (e.g., Baron & Thurston, 1973; McClelland, 1976) comparing orthographically regular pseudowords and orthographically irregular nonwords have shown that higher-order units, perhaps involving familiar combinations of letters (e.g., "th") can influence letter detection.

In many of the experiments performed in conjunction with this grant, we were concerned with frequency judgments for letters appearing in words. One of the cover tasks we studied involved a letter detection procedure of the type associated with the WSE; subjects were required to detect the presence of a U or I in a series of words and nonwords. If performance in this task were based on the formation of higher-order visual units, the formation of letter-level codes might be expected to be relatively minimal. Although it has been suggested that performance in the WSE is based on letter-level information rather than higher-order visual units (e.g., McClelland, 1976), we thought that it was important to provide a stong empirical foundation for this idea before proceeding further with the frequency judgment experiments.

To this end, Noice and Hock (1987; Section 3A) conducted a letter detection experiment, but instead of familiar, orthographically regular words, we used acronyms that were totally devoid of orthographic regularity (e.g., NBC, JFK). We found the following: 1) letters were detected more accurately in acronyms than in unfamiliar, control strings; this showed that the WSE need not be mediated by higher-order orthographic units, 2) the size of the WSE effect depended on the position of the target letter in the acronym; this showed that the effect was not due to processing the acronyms as visually familiar, global units, and 3) the advantage of the acronyms relative to the controls was significantly reduced by alternating the case of the letters within each string (e.g., nBc); this showed that the WSE for acronyms depended on the visual characteristics of the letters composing the acronyms. Our results were interpreted in conjunction with experiments by McClelland (1976), who showed that the advantage of orthographically regular pseudowords relative to nonorthographic controls was unaffected by case-alternation, and by Besner, Davelaar, Alcott, and Parry (1984), who showed that alternating the size of consecutive upper-case letters does not affect the size of the WSE. It was concluded that entries in long-term memory corresponding to familiar words and acronyms include

"recognition units" corresponding to their constituent letters. Since acronyms are experienced exclusively in upper-case, their entries in long-term memory would include representational units corresponding only to the upper-case versions of the constituent letters. Since orthographically regular words are experienced in both lower-case and upper-case formats, their entries in long-term memory would include representational units corresponding to both the upper-case and lower-case versions of the constituent letters.

F. Emphasis on Global-Level Codes Suppresses the Formation of Component-Level Codes: Evidence Obtained With Letter Strings

Although the results of the Noice and Hock (1987; Section 3A) experiment indicated that lexical entries in long-term memory for familiar words include codes corresponding to their constituent letters, it is clear that other information is also stored in the lexical entry for a word, for example, its pronunciation. Evidence for the latter has been obtained in experiments concerned with the processing of irregular words; since the way these words are pronounced does not comply with normal pronunciation rules, their pronunciation must be retrieved from their lexical entries in long-term memory (e.g., Stanovich & Bauer, 1978).

In an earlier experiment, Hock, Throckmorton, Webb, and Rosenthal (1981) found that the phonological processing of words (but not of nonwords) suppressed the retention of graphemic information associated with the words. That is, previously presented nonwords that were phonologically processed were recognized more accurately when they were presented in the same case during recognition testing compared with when they were presented in a different case; this advantage of a familiar visual format was not obtained for phonemically processed words. The hypotheses tested in conjunction with our grant research (Malcus, Hock, & Cavedo, 1985, Section 2A) were as follows: 1) lexical search based on the visual information in a word would be compatible with the formation of letter-level memory units since lexical entries in long-term memory would include "recognition units" corresponding to the letters composing the word (see Sections 1E and 3A), and 2) lexical search based on a phonological representation of a word formed prior to lexical search would suppress the formation of letter-level memory units since lexical search would be based on global-level, phonological units.

The stimuli presented during this experiment were lists of orthographically regular words and nonwords. For some lists, the nonwords were orthographically regular (e.g., TIBE), for other lists the nonwords were orthographically irregular (e.g., TBEI). The reason for this contrast was that previous research by Shulman, Hornak, and Sanders (1978) suggested that lexical decisions based on a phonological representation of a word were more likely when the nonwords in the list were orthographically regular (and pronounceable) than when they were orthographically irregular (and unpronounceable). One version of each list was printed entirely in upper-case (e.g., LIST), the other in alternating-case (e.g., LiSt). The reason for introducing the alternating-case condition was to make the stimuli visually unfamiliar, further increasing the likelihood that lexical search would be based on a familiar phonemic representation of the

word than an unfamiliar graphemic representation. There were three cover tasks accompanying the presentation of the words and nonwords: 1) a Letter Detection task in which subjects were required to search throuhg each string for the presence of a "U" or "I", 2) a Lexical Decision task in which subjects were required to discriminate between the words and nonwords, and 3) an Intentional condition, in which subjects were told to remember the frequency of occurrence of the constituent letters of the strings.

Correlations were obtained between actual and estimated letter frequency for individual subjects, and were also obtained between actual frequency and the mean-estimate frequency for all the subjects in a condition. These correlations, which measured the accuracy of frequency judgments, were similar for subjects in the Intentional and Lexical Decision conditions, both of which were greater than the correlations obtained in the Letter Search condition. The most interesting results in the experiment were obtained after we "partialled out" the frequency with which the letters whose frequency was being estimated appeared in subjects' correct-recall protocols. These partial correlation coefficients measured the extent to which actual letter frequency was judged on the basis of letter-level memory units. We found that partialling-out recall-frequency reduced the size of the correlations in all conditions. Thus is, global-level memory units contributed to the judgment of the frequency of occurrence for the letters in the words and nonwords. However, the most substantial effects were obtained for the alternating-case words in the Lexical Decision condition. Correlations in this condition were reduced to the point where they were lower than the correlations obtained in the Letter Search condition, the reverse of what was obtained prior to partialling-out the contribution of global-level memory units to subjects' frequency judgments.

We concluded, therefore, that the emphasis on global-level processing in the Lexical Decision task (subjects had to decide whether or not the string was a word) contributed positively to the over-all accuracy of subjects' letter frequency judgments (relative to the letter detection task), but only so long as lexical search was visually mediated. When the likelihood of phonological mediation increased (i.e., when case-alternation rendered the words visually unfamiliar), the encoding of letter-level memory units was suppressed. It appears, therefore, that the information associated with a word that is best remembered is the information that activates the lexical entry for the word.

Consistent with the results reported by Hock, Throckmorton, Webb, and Rosenthal (1981), evidence for the suppression of letter-level memory units as a result of phonogical mediation was not obtained for letters appearing in nonwords (e.g., BLAL). Although there are no lexical representations for nonwords, partial lexical activation by nonwords was more likely to occur as a result of letter-level codes formed for the nonwords (they could share many of their letters with nonpresented words in the lexicon) compared with phonological codes formed for the nonwords (none of our nonwords were pseudohomophones). As argued above for words, the information associated with a nonword that is best remembered appears to be the information (letter-level codes) that activates entries in the lexicon.

G. Emphasis on Global-Level Codes Suppresses the Formation of Component-Level Codes: Evidence Obtained With Patterns

Experiments with patterns have provided converging evidence for the suppression of element-level codes as a result of emphasis on pattern-level coding (Hock, Smith, Escoffery, & Bates, 1985, Section 2C). Subjects in these experiments were presented a sequence of patterns comprising five circles inside a square frame with 16 possible locations (4 x 4) for the circles. Small dots were presented inside the frame in accord with the cover task used during the presentation of the patterns. One of the orienting tasks involved detecting whether or not a dot was presented inside one of the circles, another involved determining whether a dot inside one of the circles was displaced to the left or right with respect to the center of the circle. These cover tasks were contrasted with one in which subjects were instructed to remember each of the patterns. The presentation of the patterns was followed by a frequency judgment test in which subjects were required to estimate how frequently each of the 16 possible element-locations within the frame was occupied by a circle. Two of the patterns are presented below.

The results indicated that subjects in the Pattern Memory conditions estimated element-location frequency with significantly less accuracy than subjects in the element-level conditions. Furthermore, the limited ability of subjects in the Pattern Memory condition to estimate element-location frequency could be accounted for almost completely by the retrieval of pattern-level memory units. When we counted the frequency with which each location was occupied by a circle in each subject's correct-recall protocols, and computed a partial-correlation coefficient for each subject in which the recall-frequency for each location was "partialled out," the correlation between the actual and estimated frequency was reduced to nonsignficance for the subjects in the Pattern Memory condition. Since this was not the case in the element-detection conditions, it was concluded that estimates in these conditions were based on element-level memory units. Thus, Pattern Memory instructions enhanced the formation of global-level memory units (as indicated by the recall data), but suppressed the formation of element-level memory units that were the basis for relatively accurate frequency estimation.

H. The Categorization of a Pattern Does Not Depend on Element-Location Information

Although it might be claimed that the frame-relative position of the constituent elements of a pattern constitutes highly superficial information for subjects to be retaining, it could also be argued that

element-location information can be important with regard to the formation of categories comprising the patterns. Many investigators have posited that stimuli are classified as members of a category by virtue of their similarity to prototypical representations that are formed to represent the category (Posner & Keele, 1968; Reed, 1972; Homa, 1978). It is often implicit in models of this kind, particularly when they involve dot patterns, that the prototype is formed by averaging the locations of the elements of each pattern belonging to the same category. This sort of process has recently been formalized in a neural-summation model proposed by Knapp and Anderson (1984).

Hock, Tromley, and Pohlmann (1987; Section 3B) have investigated whether the category membership of a set of patterns can be based on element-location information. That is, they asked whether the memory structures for previously learned categories include "recognition units" sensitive to the locations of the constituent elements of a pattern. A special set of dot patterns was designed such that patterns belonging to the same category all had seven of eight dots in the same relative location. Although this maximized element-correspondence for members of the same category (any further element-correspondence and the patterns would have been identical), there was no evidence in subjects' similarity judgments or classification data that they encoded the location of individual elements. The representation of category structures appeared to be based on perceptual units larger than individual elements (for example, Hock, Webb, & Cavedo, 1987, Section 3C, have provided evidence that category training can increase sensitivity to orientational similarities among patterns belonging to the same category). We concluded, on the basis of these data, that element-location information is unlikely to serve as the basis for recognizing a pattern as a member of a category.

I. Encoding Superficial Details: Abstract Attributes, Not Templates

Hock, Smith, Escoffery, and Bates' (1985, Sections 1G and 2C) evidence for the encoing of element-location information therefore joins evidence based on exact repetition effects (e.g., Jacoby & Brooks, 1984; Pollatsek, Rayner, & Collins, 1984; Hock, Throckmorton, Webb, & Rosenthal, 1981; Kolers, 1976) in providing evidence for the retention of ostensively meaningless, superficial information. One way of explaining the retention of superficial details for patterns is to specify that patterns are stored and retrieved as literal, template-like pictorial copies. However, it has been argued (Neisser, 1967; Dodwell, 1970) that template models lack the flexibility required for stimulus generalization: the ability to recognize novel patterns that are altered versions of previously seen patterns (e.g., Attneave, 1957). The results reported in the series of experiments by Hock, Smith, Escoffery, and Bates indicated that superficial element-location information can be retained in the form of element-level memory units rather than template-like pictorial copies. Superficial details involving element position were abstracted from a series of patterns in the same way that semantically important, shared attributes might be abstracted from a series of patterns belonging to the same category (e.g., Posner & Keele, 1968). The results of the

experiments therefore indicate that there is no reason to assume that superficial details are abstracted and represented any differently than the attributes that are the basis for categorizing previously seen stimuli and novel stimuli.

J. Encoding Superficial Details: Evidence For Pattern-Analyzing Memory

If some of the patterns presented by Hock, Smith, Escoffery, and Bates (1985, Sections 1G and 2C) were shifted to the left, and others shifted upward, there would be no change in the global-level content of the series (the patterns would remain the same). Nonetheless, the frequency with which circles appeared in various locations within the frame would change. Of what use, then, is the superficial information abstracted from sequences of patterns?

The experiments conducted with patterns by Hock, Smith, Escoffery, and Bates, and with words by Hock, Malcus, and Hasher, 1986, Section 2A), suggest that a relatively small sample of stimuli (e.g., 34 patterns in one set of experiments), is sufficient for subjects to abstract superficial details that characterize the particular set of materials that they are processing. Although this information will not, in general, contribute to the categorization of individual stimuli, it has the potential to facilitate further processing of materials possessing the same superficial characteristics. This is the sort of information that Kolers (1976) appears to have had in mind in his studies of pattern-analyzing memory.

An experiment which could test the hypothesis that repetition effects of the kind that Kolers (1976) attributes to pattern-analyzing memory can be based on the frequency of occurrence of superficial characteristics of the stimuli might proceed as follows. The initial phase of the experiment would involve the presentation of a series of patterns (comprising five circles) with a particular distribution of element-location frequencies. Subjects would be required to determine whether or not a dot is present inside one of the circles. They would then be transferred to another series composed of the same patterns, but displaced to new locations within the frame. In one condition, the displacement would preserve the original distribution of element-location frequencies (for example, a pattern that had an element in the upper-left corner of the frame might be shifted away from this location and replaced by a different pattern with an element in the upper-left corner). In a second condition, the displacement of the original patterns would result in a different distribution of element-location frequencies. Subjects in the transfer phase would again be required to determine whether or not a dot was present inside one of the circles (the target-dot would be located in different circles compared with the initial phase of the experiment). If our pattern-analizing hypothesis is correct, dot-detection time would be faster in the condition with the repeated frequency distribution (and there will be no difference in pattern recall). Further support for the pattern-analyzing hypothesis could then by sought in a second experiment in which subjects task in the transfer phase will involve higher-order attributes of the patterns (e.g., whether or not they include diagonally arranged circles). We would again expect performance in the Phase 2, diagonal-detection task to benefit from

the repetition of the superficial frequency distribution acquired with the Phase 1, dot-detection task.

K. Encoding Superficial Details Affects Identification More Than Categorization

Hock, Rosenthal, and Stenquist (1985; Section 3D) performed a series of experiments based on the category effect in visual search. The category effect refers to evidence that it is easier to detect the presence of a target digit in a field of letters than it is to detect the presence of a target letter in a field of other letters. Hock, Rosenthal, and Stenquist's study focussed on a catch-trial presented at the conclusion of a series of between-category trials (looking for digits among letters). On these catch-trials, digits other than the expected target-digit were presented. With relatively little practice (96 trials), subjects were likely to respond slowly and incorrectly on these trials, the effect diminishing with increased practice until performance on the catch-trials was no worse than that observed on standard target-absent trials (after 384 practice trials). These results were consistent with the view that superficial attributes that are abstracted from relatively short series of stimuli can facilitate the processing of information specific to individual category members to a greater extent than information which is shared by members of the same category.

L. Coding Spatial Relations Facilitates Judgments of Element Position

In our initial discussion of the series of experiments involving frequency judgments for element-location in Section 1G (Hock, Smith, Escoffery, & Bates, 1985, Section 2C), we distinguished between the effects of element-level and global-level codes on judgments of element-location frequency. Since global-level pattern codes must include information concerning the relative location of elements in the pattern, the distinction between element-level and global-level codes reduces to a distinction between frame-relative and element-relative position codes.

Although the experiments indicated that judgments of element-location frequency were relatively inaccurate as a result of global (pattern)-level processing, further evidence, accumulated over four experiments, produced surprising evidence that the formation of pattern-level memory units was associated with more precise encoding of element position than the formation of element-level memory units. Since the formation of pattern-level memory units requires that the position of elements be coded relative to other elements in the pattern, pattern-level memory units may result in relatively precise judgments of element-position because of constraints arising from the encoding of multiple spatial relations among the elements of the pattern (e.g., an element might be alongside one element, diagonally below a second element, and relatively far from a third element directly above it). Despite this precision, the relational information in pattern-level memory units can be useless for estimating element-location frequency if the global locations of the patterns, relative to the frame, have not also been encoded (pattern-recall data in the Hock, Smith, Escoffery, and Bates

experiments indicated that frame-relative coding of pattern position was imprecise).

Evidence of a much different kind has been obtained which converges with the conclusion that the encoding of spatial relations can facilitate the accuracy with which subjects judge an element's position. This evidence came from experimental work done with Christina's World, a well-known painting of Andrew Wyeth's (Hock, 1984; Section 3E). The painting is set on an open field, with an uphill slope that terminates as a horizon line against the sky. There are three important objects in this setting: a barn and a farmhouse along the horizon, and Christina, who is lying in the field in the foreground of the painting. Although Christina's face is hidden, the orientation of her body creates a strong spatial alignment between her and the farmhouse. The perception of this alignment was crucial for subjects' performance. After examining a copy of the painting, subjects were presented five alternative versions in a subsequent recognition test. The five alternatives varied with regard to the position of the barn along the horizon. Hock found that subjects could recognize the correct location of the barn at a better than chance rate only when the spatial alignment between Cristina and the farmhouse was perceived. Recognition performance was at chance when the painting was presented upside-down during its initial presentation, and was also at chance when Christina's body was shielded from the subjects' view, leaving vision of only the back of her head. We could conclude that judments of the barn's position were enhanced by the constraints imposed by encoded relations among the objects in the painting; the perceived alignment between Christina and the farmhouse appears to have provided an essential constraint on the relative positions of objects within the scene.

M. Subjects Retain Information Involving Spatial Relations

The analyses described in the previous section suggest that the experiments performed by Hock, Smith, Escoffery, and Bates (1985, Sections 1G and 2C) may have underestimated how well subjects can remember relations among elements. Evidence supporting this conclusion comes from a series of experiments performed by Rose and Hasher (Section 2G). In these experiments, subjects were presented simple patterns comprising pairs of elements in one of four arrangements: one above the other, one alongside the other, one to the left and above the other, and one to the right and above the other. Pairs of these patterns were presented sequentially. Sometimes the pattern remained the same on both presentations, sometimes it changed. When the pattern changed, it could change with respect to: 1) the global location of the pattern on the screen, 2) the distance between the elements, and 3) the relation between the elements. Rose and Hasher found that changes in spatial relationships were recognized more accurately than changes in the distance between the elements or changes in the global location of the pattern. The results obtained for changes in distance were surprising in that contractions in distance were more accurately recognized than expansions in distance when the elements were relatively close together, and vice versa when the elements were relatively far apart. Only the results for the close elements were consistent with Weber's Law. All these results

were unaffected by variations in the time interval between the two patterns or the presence of a fixation point.

N. Spatial Relations Between Elements Are Coded Explicitly

Our ability to make judgments of spatial relations may or may not mean that spatial relations are actually encoded in memory. For example, being able to recall that one element or object was to the left of another does not mean that the relation "to the left (or right)" was explicitly encoded in memory. It is possible that in the initial perception of the two elements (or objects), the memory codes formed were based on a Cartesian reference frame. That is, the frame-relative position was independently coded for each element. Since the left/right relation of the two elements is implicit in these codes, subjects' recollection of one element being to the left of the other could be derived from the Cartesion rather than relational memory codes.

Rose and Hasher (Section 2G) have used the frequency discrimination technique to provide evidence that spatial relations are encoded explicitly. They devised a set of nine context or background scenes, each containing a large object (e.g., a crib, a picnic table, etc.). Into each background scene (defined by the large object) was inserted a second object that could occur in one of several spatial relationships with the context (background) object. Included among the relationships were in front, above, inside, etc. Across the various contexts, the frequency with which a relationship occurred was varied. For example, the subject may have seen four different objects inside the crib and two different objects in front of the picnic table. Rose and Hasher found that subjects could accurately judge the relative freqency with which each relation occurred in conjunction with a particular context (background) object. This was the case for a variety of instructional conditions, including instructions that were completly incidental to the frequency judgment test.

In an additional experiment, scenes were selected such that each pair of objects appeared in above/below, side-by-side, and diagonal relations, but with different frequency (e.g., once in above/below, three times side-by-side). Following the presentation of a series of such scenes, subjects could accurately judge the frequency with which each pair appeared in each relation. Taken together, these results suggested that frequency judgments were based, not on Cartesian codes for the location of individual objects, but on explicity coded spatial relations.

An experiment that could provide an additional test of the view that spatial relations are coded explicitly would focus entirely on the spatial relation of diagonal-alignment for pairs of elements (empty circles); one circle would be below and to the right of the second circle. This relation of two circles would reappear, with varying frequency, in multiple locations throughout a 4x4 matrix of possible circle locations. The diagonal would be embedded in a more complex pattern of circles in one of two ways: for the patterns belonging to Set A, it would be likely that the diagonal relation would be explicitly coded as part of each pattern's description; for the patterns belonging to Set B, it would be unlikely that the

diagonal relation would be coded (a pretest would be necessary in order to confirm that the encoding oBf the diagonal relation is more likely for the Type A than the Type B patterns). Following the presentation of these patterns (Set A to one group of subjects, Set B to another), subjects would be required to judge the frequency with which the diagonal arrangement of two circles appeared in different locations within the 4x4 matrix. Since the two sets would be matched with regard to the Cartesian coordinates of the elements forming the diagonal, no difference between Sets A and B in the accuracy of frequency judgments would be expected if the diagonal relation was encoded implicitly (i.e., if the position codes were entirely Cartesian). However, if the the diagonal relation was coded explicitly for Set A, but not for Set B, then frequency judgments would be more accurate for Set A than Set B.

O. Encoding Spatial Relations Facilitates Object Identification

It was concluded in Section 1L that the encoding of spatial relations can facilitate the accuracy with which subjects judge an object's position. In a further experiment, Rose and Hasher (Section 2G) showed that the encoding of spatial relations can facilitate the speed with which subjects identify objects entering into the spatial relations. Subjects in this experiment were presented a series of two-object scenes with one of four relations among the objects (above/below, side-by-side, diagonal, and front/behind). They were then given a recognition test in which they were required to indicate whether the presented objects were the same as those presented earlier. When the objects were the same, they could appear in the same or different spatial relationship. Subjects were faster when the relationship was the same. This was the case regardless of whether the test scenes were preceded by a scene constituting part of the test scene (the priming condition) or part of another scene (the mispriming condition).

P. The Effect of Pattern Structure on Frequency Judgments

Patterns with simple, symmetrical structures are easier to identify (Clement, 1964) and reproduce (Bell & Handel, 1976) than patterns with more complex, asymmetrical stuctures. An experiment by Hock, Bates and Field (Section 2H) examined whether similar differences would be observed when subjects judged the frequency of occurrence of simple and complex patterns. Patterns comprising five circles (within an imaginary 3 x 3 matrix) were presented within a 4 x 4 frame with 16 possible locations for the circles. The patterns were presented either four or eight times, sometimes at the exact same location within the frame, sometimes in different quadrants within the frame (this ultimately had no effect on performance). Subjects participated in one of two tasks during the initial presentation of the patterns. One group searched for the presence of a dot within one of the circles, the other was told to try to remember the patterns. This was followed by a frequency estimation test during which patterns were presented without the surrounding frame. Subjects did a better job of discriminating between the low and high frequency patterns in the Pattern memory compared with the Dot Detection condition. They

also discriminated between high and low frequency "simple" patterns to a greater extent than they discriminated between high and low frequency "complex" patterns. This interaction, however, fell short of statistical significance at the .05 level.

Q. Theoretical Conclusions

Frequency judgments for words, patterns, and object-relations depend on backward searches through memory that re-activate previously experienced versions of the words, patterns, and object-relations that are stored in episodic memory (Tulving, 1972). These global-level memory units enable people to make important decisions about such significant problems as who to model one's behavior after (Perry & Bussey, 1979), whether or not to believe an assertion is actually true (Hasher, Goldstein, & Toppino, T., 1977), and what is going to happen next.

Frequency judgments for the constituent elements of words and patterns depend on the introduction of orienting tasks which direct attention to the constituent elements of the word or pattern; orienting tasks that direct attention to global-level units can suppress the formation of component-level memory units. These component-level memory units may be the basis for pattern-analyzing operations (Kolers, 1976) that provide for "fluency" in reading and pattern identification.

References

Attneave, F. (1957). Transfer of experience with a class-schema to identification-learning of patterns and shapes. *Journal of Experimental Psychology, 54,* 81-88.

Baron, J., & Thurston, I. (1973). An analysis of the word-superiority effect. *Cognitive Psychology, 4,* 207-228.

Bell, H. H., Handel, S. (1976). The role of pattern goodness in the reproduction of backward masked patterns. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 139-150.

Besner, D., Davelaar, E., Alcott, D., & Parry, P. (1984). Wholistic reading of alphabetic print: Evidence from the FDM and the FBI. In L. Henderson (Ed.), *Orthographies and reading: Perspectives from cognitive psychology, neuropsychology, and linguistics.* Hillsdale, N.J.: Erlbaum Associates.

Clement, D. E. (1964). Uncertainty and latency of verbal naming responses as correlates of pattern goodness. *Journal of Verbal Learning and Verbal Behavior, 3,* 150-157.

Dodwell, P. C. (1970). *Visual pattern recognition.* New York: Holt, Rinehart and Winston.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87,* 272-300.

Fisk, A. (1986). *American Psychologist, 41,* 215-216.

Fisk, A. D., & Schneider, W. (1983). Category and word search: Generalizing search principles to complex processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 9,* 177-195.

Greene, R. L. (1984). Incidental learning of event frequency. *Memory & Cognition, 12,* 90-95.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16,* 107-112.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356-388.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case for frequency of occurrence. *American Psychologist, 39,* 1372-1388.

Hasher, L, Zacks, R. T., Rose, K. C., & Sanft, M. (1987). Truly incidental incoding of frequency information. *American Journal of Psychology,* in press.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 88, 297-306.

Hock, H. S. (1984). Christina's World: Imaginary perspectives and the encoding of spatial alignment relations. In W. R. Crozier and A. J. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North-Holland.

Hock, H. S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 12, 232-240.

Hock, H. S., Rosenthal, A., & Stenquist, P. (1985). The category effect in visual search: Practice effects on catch-trials. *Perception & Psychophysics*, 37, 73-80.

Hock, H. S., Throckmorton, B., Webb, E., & Rosenthal, A. (1981). The effects of graphemic and phonemic processing on the retention of words and nonwords. *Memory & Cognition*, 9, 461-471.

Hock, H. S., Smith, L., Escoffery, L., & Bates, A. (1985). Pattern perception: Encoding element location. Paper read at the 26th Annual Meeting of the Psychonomic Society, Boston, MA.

Hock, H. S., Tromley, C., & Pohlmann, L. (1987). Perceptual units in the acquisition of visual categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.

Hock, H. S., Webb, E., & Cavedo, L. C. (submitted). Perceptual learning in visual category acquisition.

Homa, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 407-416.

Jacoby, L. L, & Brooks, L. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 18). New York: Academic Press.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 616-637.

Kolers, P. A. (1976). Pattern-analyzing memory. *Science*, 191, 1280-1281.

Malcus, L., Hock, H.S., Cavedo, L. C., Smith, L. B. (1985). Orthographic processing of strings affects frequency discrimination for component letters. Paper presented at meeting of Eastern Psychological Association, Boston, Mass.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 80-91.

Naveh-Benjamin, M., & Jonides, J. (1986). On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 378-386.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Noice, H., & Hock, H. S. (1987). A word superiority effect with non-orthographic acronyms: Testing for unitized visual codes. Paper to be presented at the annual meeting of Eastern Psychological Association.

Perry, D. G. & Bussey, K. (1979). The social learning theory of sex differences: Imitation is alive and well. *Journal of Personality and Social Psychology*, 37, 1699-1712.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, 113, 426-442.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353-363.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382-407.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus materials. *Journal of Experimental Psychology*, 81, 275-280.

Shulman, K. G., Hornak, R., & Sanders, S. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition*, 6, 115-123.

Stanovich, K.E., & Bauer, K. W. (1978). Experiments on the spelling-to-sound regularity effect in word recognition. *Memory & Cognition*, 6, 410-415.

Tulving, E. Episodic and semantic memory (1972). In E. Tulving & W. Donaldson (Ed.), *Organization and memory*. New York: Academic Press.

Tversky, A., & Kahnemen, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Tversky, B., Hemenway, K. (1984). Objects, parts and categories. *Journal of Experimental Psychology: General*, 113, 169-193.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Zacks, R. T., Hasher, L., Hock, H. S. (1986). Inevitability and automaticity: A response to Fisk. *American Psychologist*, 41, 216-218.

Frequency Discrimination: Assessing Global-Level and Element-Level Units in Memory

Howard S. Hock and Lawrence Malcus
Florida Atlantic University

Lynn Hasher
Temple University

Subjects' knowledge of how often various events occur was used to assess the retention of memory units for word-like strings of letters. A series of strings was presented at one of three exposure durations. Within the series, the frequencies of occurrence of different strings and of the letters composing the strings were varied orthogonally. At relatively long exposure durations, subjects could discriminate the frequency of occurrence for both strings and their constituent letters. The formation of global-level (string) memory units was indicated by judgments of string frequency being unaffected by either the frequencies of their component letters or experimental conditions (brief exposures) that prohibited accurate judgment of letter frequency. Although judgments of letter frequency were sometimes biased by the frequency of the strings containing the letters, the success with which the judgments discriminated different levels of letter frequency did not depend on the activation of string-level memory units. Furthermore, subjects' frequency judgments for letters were not predictable from their recall of the strings containing the letters. These results, which could not be explained by Tversky and Kahneman's (1973) "availability heuristic," provided evidence for the formation of element-level (letter) memory units. A converging experiment established that element-level frequency information could be abstracted from words as well as nonwords, and further, that this information was stored in long-term memory.

Of central concern to both perceptual and cognitive theories of visual processing (e.g., Hochberg, 1981; Neisser, 1967) is the issue of whether the functional units are elements, subsets of elements, or the entire visual array. For example, the functional units in the identification of printed words could be individual letters, orthographically regular combinations of letters, or the entire word. Regardless of the size of functional units during the identification of printed strings of letters, our concern in this article was to determine whether informational units of different size are stored in memory.

It is well established that people are sensitive to information about the frequency with which events occur (Hintzman & Block, 1971; Underwood, 1969). This information appears to be processed with little effort (Hasher & Zacks, 1979) or intention (Howell, 1973). Our experiments were designed to capitalize on this sensitivity to occurrence-rate information as a way of identifying memory units for words and word-like items. We did this in our first experiment by varying the frequency of occurrence of strings of letters orthogonally to the frequency of occurrence of the individual letters composing the strings. After presenting the strings at varying exposure durations, we asked subjects to judge either letter or string frequency. Our results provided con-

verging evidence for the retention of both global-level (string) and element-level (letter) memory units. Retrieval of memory units at both levels influenced subjects' judgments of letter frequency. String frequency judgments appeared to be influenced only by string-level memory units.

Previous researchers have been concerned with frequency judgments for items that were physically identical for each repetition in a list as well as information abstracted from the items in a list. Gude and Zechmeister (1975) and Burnett and Stevenson (1979) have compared frequency judgments for sentences that were literally identical on each repetition with sentences that were literally different but kept the same meaning on each repetition. Jacoby (1972) and Rowe (1973a, 1973b) have similarly compared frequency judgments for words that were literally identical on each repetition with frequency judgments for words that were literally identical but varied in meaning (homonyms) and words that were literally different but the same in meaning (synonyms). Our approach is somewhat different. We are interested in frequency judgments for information abstracted from the items in the list, but the abstraction of interest is across the list rather than within the individual items composing the list. From this point of view, letter-level memory units, coded for frequency of occurrence, could constitute an abstract description of the compositional characteristics of the list. By analogy with category acquisition research, the stimulus list presented at some point in time could be thought of as a particular information category, with the global-level (string) units corresponding to the exemplars of the category and the element-level (letter) units corresponding to an abstract, featural description of the category.

The initial purpose of Experiment 1 was to determine if, at relatively long presentation durations, subjects' judgments would discriminate different levels of occurrence frequency for both

Correspondence concerning this article should be addressed to Howard S. Hock, Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431.

233   H. HOCK, L. MALCUS, AND L. HASHER

the strings and the letters composing the strings. Correlational data suggest that people have reliable knowledge about the frequency with which both words (Carroll, 1971) and individual letters (Attneave, 1953) occur in natural language. It was important to demonstrate that under relatively ideal circumstances, subjects could discriminate frequency of occurrence at both the element (letter) and global (string) levels. By varying letter and string frequency orthogonally, we could determine whether frequency judgments at the element level were derived from stored frequency information at the global level and vice versa.

Experiment 1 also manipulated the presentation duration for the strings, the purpose being to determine the levels of memory representation, if any, that are sacrificed when processing constraints are increased (here, by reducing the presentation duration). If, for example, brief exposure durations eliminated subjects' ability to judge the frequency of letters while leaving intact their ability to judge the frequency of strings, it would provide evidence that judgments of frequency for strings were not derived from stored frequency information for their constituent letters.

The final purpose of the first experiment was to determine whether subjects use a strategy in which letter-frequency judgments are based on the activation of string-level memory units. According to one such strategy, subjects' estimates of letter frequency would increase if they activate the memory representation for a string containing the letter. This strategy, which is similar to the "availability heuristic" proposed by Tversky and Kahneman (1973), could be the basis for accurate judgments of letter frequency. The latter have argued that estimates of the frequency of various events depend on the activation in memory of specific instantiations or associates of the events. The activation of strings could be responsible for the accurate judgment of letter frequency because in our experiments, as in natural language, the reason some letters are higher in frequency than others is that high-frequency letters occur in more different strings than do low-frequency letters. As a result, there is more opportunity for subjects to activate memory representations of strings containing high-frequency letters compared with strings containing low-frequency letters, and further, to use this as the basis for judgments of the frequency of occurrence of constituent letters.

Experiment 1 was designed to control for the possibility that subjects would be biased to judge letters as high in frequency simply because they occurred in "more available," high-frequency strings. The orthogonal manipulation of letter and string frequency prevented such a bias from being the basis for subjects' discrimination between high- and low-frequency letters. Subjects' free-recall protocols were also used to investigate the role of string activation in letter-frequency judgments. In analyzing the recall data we assumed that the likelihood of a string being recalled was directly related to the likelihood that its memory representation was activated while subjects were making judgments of letter frequency.

Experiment 1

Method

Subjects. Seventy-two students in undergraduate psychology classes at Florida Atlantic University voluntarily participated in this experiment without pay.

Table 1

The Consonants and Strings Assigned to the Four Experimental Conditions Used in Experiments 1 and 1A

| Consonant frequency | String frequency | | | |
| | High (6) | | Low (3) | |
	Consonants	Strings	Consonants	Strings
High (12)	N,P,W	AWUP WENY NIPO	S,B,M	SUMY EMSA BOIM MUBY ASEB OBIS
Low (6)	X,H,G,C,R,F	ROCY AGIF HUXE	Z,T,K,L,V,D	ZAKY TIDU OLEV UZIL KEDO VATY

Design. The experiment was conducted in three phases. Phase 1 involved presenting a sequence of 72 letter strings at one of three exposure durations. In Phase 2, independent groups of subjects judged the frequency of occurrence for either the strings or the letters composing them. Only subjects judging letter frequency during Phase 2 participated in Phase 3, which involved recalling the strings presented during Phase 1. The orthogonal combination of the three exposure durations and the two frequency judgment conditions (letter vs. string) generated six between-subject experimental conditions. Twelve subjects were randomly assigned to each condition. The stimuli, which are presented in Table 1, were designed such that the orthogonal combination of two variables, high- versus low-frequency letters and high- versus low-frequency strings, produced four within-subject conditions.

Stimuli. The initial step in generating the stimuli was to assign three consonants to each of the two high-letter-frequency conditions and six consonants to each of the two low-letter-frequency conditions. The average frequency of usage in English (Mayzner & Tresselt, 1965) was virtually identical for the four sets of consonants. The five vowels, plus Y, were then combined with the consonants to produce the 18 orthographically regular four-letter strings presented in Table 1.

In the high-letter-frequency, high-string-frequency condition, each string was presented a total of six times. Because each consonant in this condition appeared in two different strings, each was presented a total of 12 times. In the low-letter-frequency, high-string-frequency condition, each string was again presented six times, but now each consonant appeared in only one string. As a result, each was presented a total of six times. The same logic was applied in the two remaining stimulus conditions. Following this procedure, 18 strings were constructed. Six were presented six times each and 12 were presented three times each, producing a total of 72 stimuli. As can be seen in Table 1, consonants that appeared in two different strings were presented in different positions in each string and in combination with different vowels and consonants. Each vowel, including Y, appeared equally often; two vowels were always combined with two consonants. The strings were all typed in upper-case, Letter Gothic type font.

None of the strings used in the experiment were English words. However, most could be converted to words by changing the identity of a single letter, leaving position unchanged (e.g., WENY could become WENT). The strings assigned to the four experimental conditions were matched in that two thirds of the items in each condition could be converted to English words by changing one letter.

Procedure. The 72 strings were presented, in random order, at one of three presentation rates. The fastest rate, which was achieved by holding

Table 2
Judgments of Letter Frequency: The Proportion of "12" Responses

Exposure duration (s)	Frequency of letters in high-frequency strings				Frequency of letters in low-frequency strings				Letter frequency for all strings			
	12	6	12 − 6	M	12	6	12 − 6	M	12	6	12 − 6	M
					Experiment 1							
4.2	.70	.26	.44	.48	.44	.22	.22	.33	.57	.24	.33	.41
1.2	.36	.17	.19	.27	.39	.24	.15	.32	.38	.21	.17	.30
0.2	.42	.53	−.11	.48	.40	.25	.15	.33	.41	.39	.02	.40
M	.49	.32	.17	.41	.41	.24	.17	.33	.45	.28	.17	.37
					Experiment 1A							
0.2	.48	.47	.01	.48	.44	.42	.02	.43	.46	.45	.01	.46

Note. The proportion of "12" responses refers to how often subjects judge a letter to have occurred 12 rather than 6 times.

the advance button on the Kodak Carousel projector in a depressed position, resulted in an exposure duration of approximately 0.2 s per string and a duration of approximately 0.8 s between exposures. The slower two presentation rates resulted in exposure durations of approximately 1.2 and 4.2 s per string, the duration between exposures remaining at approximately 0.8 s. The projected width of each string was approximately 15 cm. Each letter was approximately 2.2 cm wide. Because the experiment was conducted in classrooms, the visual angle intercepted by each string varied from subject to subject but was never greater than 3.0°. Prior to the presentation of the strings, each group of subjects was instructed to try to remember the information presented on the screen. There were no directions concerning whether subjects should attend to individual letters or strings and no indication that we would subsequently assess frequency of occurrence information.

Following the presentation of the strings, subjects in alternate seats were assigned to two groups. One group made decisions about consonant frequency, the other group about string frequency. Booklets were prepared with one consonant (or string) per page, each booklet comprising a different random order. Subjects judging string frequency were required to circle either the "3" or "6" typed above each string. Subjects judging letter frequency were required to circle either the "6" or "12" typed above each consonant. Following these responses, subjects in the letter (consonant) judgment condition were instructed to recall as many strings as possible.

Results

Letter-frequency judgments. The overall percentages of letter-frequency judgments that were correct were 67%, 59%, and 52%, in the 4.2-, 1.2-, and 0.2-s conditions, respectively. These data indicated that subjects could discriminate letter frequency for the 4.2-s and 1.2-s exposures (chance was 50% correct), but frequency discrimination following the 0.2-s exposures was equivocal. Average performance in the latter condition was close to chance (52% correct), with subjects performing below chance (45%) for letters from high-frequency strings and above chance (58%) for letters from low-frequency strings.

For purposes of analysis, the dependent variable was the proportion of subjects' responses for which letters were judged to have occurred 12 times (half the letters were presented 12 times, half 6 times). This measure allowed us to determine whether letter-frequency judgments (a) discriminated between high- and low-frequency letters (as indicated by positive 12–6 difference

scores), and (b) were biased by the frequency of the strings containing the letters (as indicated by the mean proportion of "12" judgments at each level of string frequency). These results are summarized in Table 2. In addition to showing the previously described effect of exposure duration on discrimination accuracy, the data in Table 2 indicate a general bias to judge letters as low in frequency.

An analysis of variance (ANOVA) on the proportion of "12" responses in subjects' letter judgments indicated that there was a significant interaction between the effects of letter frequency and exposure duration, $F(2, 33) = 6.20$, $p < .01$, $MS_e = .045$. Tests of simple effects indicated that the effect of letter frequency was significant for the 4.2-s and 1.2-s exposure durations, $F(1, 33) = 28.32$, $F(1, 33) = 24.08$, $p < .01$, $MS_e = .045$, respectively, but was not significant for the 0.2-s exposure duration, $F(1, 33) < 1.0$, $MS_e = .045$. High- and low-frequency letters were not differentiated more accurately when they were presented in high-frequency compared with low-frequency strings. That is, the interaction between letter frequency and string frequency was not significant, $F(1, 33) < 1.0$, $MS_e = .086$. The three-way interaction between exposure duration, letter frequency, and string frequency also was not significant, $F(2, 33) = 2.00$, $p > .05$, $MS_e = .086$.

With regard to response bias, the main effect of string frequency, $F(1, 33) = 5.63$, $p < .05$, $MS_e = .042$, and the interaction between string frequency and exposure duration, $F(2, 33) = 3.63$, $p < .05$, $MS_e = .042$, were significant. The latter two effects were obtained because frequency judgments were biased to be relatively high for letters from high-frequency strings, at least for the 4.2-s and 0.2-s exposure durations. This bias, which was smaller and unreliable in the replication reported in Experiment 1A, did not directly contribute to the discrimination of letter frequency because string frequency and letter frequency were varied orthogonally in both experiments.

A further analysis, in which the letters used in the experiment replaced subjects as the random variable in the ANOVA, indicated that the results were generalizable over the letters used in the experiment.[1] The interaction between letter frequency and ex-

[1] Because the experiment was designed with different numbers of letters assigned to the high- and low-frequency conditions, we performed a least-

Table 3
Judgments of String Frequency: The Proportion of "6" Responses

Exposure duration (s)	Frequency of strings with high-frequency letters				Frequency of strings with low-frequency letters				String frequency for all letters			
	6	3	6 – 3	M	6	3	6 – 3	M	12	6	6 – 3	M
					Experiment 1							
4.2	.75	.26	.49	.51	.81	.14	.67	.48	.78	.20	.58	.49
1.2	.81	.29	.52	.55	.81	.38	.43	.60	.81	.34	.47	.58
0.2	.67	.43	.24	.55	.84	.36	.48	.60	.76	.40	.36	.58
M	.74	.33	.41	.54	.82	.29	.53	.56	.78	.31	.47	.55
					Experiment 1A							
0.2	.67	.37	.30	.52	.79	.46	.33	.63	.73	.42	.31	.58

Note. The proportion of "6" responses refers to how often subjects judge a letter to have occurred 6 rather than 3 times.

posure duration was again significant. $F(2, 28) = 8.21$, $p < .005$, $MS_e = .011$. Tests of simple effects again indicated that the effect of letter frequency was significant at the 4.2-s and 1.2-s exposure durations, $F(1, 28) = 39.56$, $F(1, 28) = 11.24$, $p < .005$, $MS_e = .011$, respectively, but not the 0.2-s exposure duration. $F(1, 42) < 1.0$, $MS_e = 0.11$. Once again, the interaction between letter frequency and string frequency was not significant. $F(1, 42) < 1.0$, $MS_e = .020$, but the main effect of string frequency, $F(1, 28) = 4.39$, $p < .05$, $MS_e = .020$, and the interaction between exposure duration and string frequency, $F(2, 28) = 6.38$, $p < .01$. $MS_e = .011$, were significant. The one difference from the analysis in which subjects was the random variable was that the three-way interaction between exposure duration, letter frequency, and string frequency was now significant. $F(2, 28) = 5.50$, $p < .05$. $MS_e = .011$. This interaction, which was not reliable over the full set of participating subjects, reflected the relatively small effect of exposure duration on frequency discriminability for letters from low-frequency strings. This insensitivity to exposure duration may have been the result of performance being relatively poor (at "floor") for the letters from low-frequency strings. Our clearest evidence for successful letter-level frequency discrimination (and its elimination at brief exposure durations) was obtained for letters from high-frequency strings.

The above analyses indicated that subjects could discriminate high-frequency from low-frequency letters for 1.2- and 4.2-s exposure durations. It remained possible, however, that their judgments of letter frequency were based on the activation of global-level memory units (i.e., strings) containing the letters rather than the retrieval of letter-level frequency information. To evaluate this possibility, letter-frequency judgments were compared at the 1.2-s and 4.2-s exposure durations for high-frequency consonants embedded in low-frequency strings and low-frequency consonants embedded in high-frequency strings (these data are

part of the full data set presented in Table 2). If letter-frequency judgments were based on the activation of strings containing either high- or low-frequency consonants, differences in string frequency would have favored the activation of strings containing low-frequency letters. Nonetheless, high-frequency letters were judged as higher in frequency than low-frequency letters for both the 1.2-s exposure duration, $t(11) = 3.74$, $p < .005$, and for the 4.2-s exposure duration, $t(11) = 2.57$, $p < .05$.

String-frequency judgments. The overall percentages of string frequency judgments that were correct were 80%, 79%. and 68%, in the 4.2-, 1.2-, and 0.2-s conditions, respectively. The proportion of responses for which the strings were judged to have occurred six times (half the strings had been presented six times. half three times) are presented in Table 3. It can be seen from the 6–3 difference scores that subjects' judgments discriminated between the two levels of string frequency at all three exposure durations. It can also be seen from the mean proportion of "6" responses in Table 3 that differences in letter frequency introduced little response bias into subjects' judgments of string frequency.

An ANOVA was performed in which frequency judgments were contrasted for high- and low-frequency strings. which in turn were composed of high- and low-frequency letters. The analysis indicated that the effect of string frequency on the proportion of "6" responses in subjects' string judgments was significant. $F(1, 33) = 107.07$, $p < .001$. $MS_e = .074$. This effect was obtained for all three exposure durations; the interaction between exposure duration and string frequency was not significant. $F(2, 33) = 1.96$, $p > .05$, $MS_e = .074$. Although string-frequency judgments were slightly higher for the low-letter-frequency than the high-letter-frequency condition. the main effect of letter frequency was not significant. $F(1, 33) < 1.0$, $MS_e = .035$. Two interactions were marginally significant. The three-way interaction between string frequency, letter frequency, and exposure duration. $F(2, 33) = 3.30$, $MS_e = .026$, barely reached significance at the .05 level. The interaction between string frequency and letter frequency, $F(1, 33) = 4.09$, $MS_e = .026$, fell just short of significance at the .05 level. These interactions may have been due to uncontrolled characteristics of some of our items (e.g., certain items may have been more likely than others to remind subjects of familiar words). This was suggested by the results of an additional ANOVA in which items replaced subjects as the random variable.

squares analysis of these data. A different subset of letters was assigned to the four stimulus conditions generated by the orthogonal combination of letter frequency and string frequency. The latter were therefore treated as between factors in the analysis. Because each letter was presented at all three exposure durations, exposure duration was treated as a within factor in the analysis. The same approach was taken for the other item analyses reported in this experiment.

Table 4
Probability of a Letter Being Present in at Least One String Recalled by a Subject (Experiment 1)

Exposure duration (s)	Frequency of letters in high-frequency strings				Frequency of letters in low-frequency strings				Letter frequency for all strings			
	12	6	12 – 6	M	12	6	12 – 6	M	12	6	12 – 6	M
4.2	.81	.53	.28	.67	.64	.40	.24	.52	.73	.47	.26	.60
1.2	.50	.47	.03	.49	.58	.14	.44	.36	.54	.31	.23	.43
0.2	.61	.31	.30	.46	.14	0	.14	.07	.38	.16	.22	.27
M	.64	.44	.20	.54	.45	.18	.27	.32	.55	.31	.24	.43

The results of this item analysis again indicated that the effect of string frequency on the proportion of "6" responses was significant, $F(1, 14) = 79.55$, $p < .001$, $MS_e = .033$, and the interaction between string frequency and exposure duration was not significant, $F(2, 28) = 2.09$, $p > .05$, $MS_e = .026$. However, neither the interaction between string and letter frequency, $F(1, 14) = 1.22$, $p > .05$, $MS_e = .033$, nor the three-way interaction between string frequency, letter frequency, and exposure duration, $F(2, 28) = 1.22$, $p > .05$, $MS_e = .026$, was significantly generalizable over the items used in the experiment. The interaction effects obtained with subjects serving as the random variable were thus limited to a relatively small number of strings in the stimulus list.

The results obtained for the 0.2-s presentation duration indicated that string frequency could be discriminated under conditions that prohibited the retention of letter frequency information (subjects' frequency judgments following the 0.2-s exposures did not discriminate between the two levels of letter frequency). This evidence that frequency judgments for strings were not based on the frequency of occurrence of their constituent letters was supplemented by the following. At each exposure duration, string-frequency judgments were compared for high-frequency strings composed of low-frequency letters and low-frequency strings composed of high-frequency letters. If string-frequency judgments were based on the frequency of occurrence of the letters composing each string, differences in letter frequency would have favored judging the low-frequency strings as high in frequency (the relevant data are part of the full data set presented in Table 3). Nonetheless, high-frequency strings were judged as higher in frequency than low-frequency strings at the 0.2-s exposure duration, $t(11) = 6.87$, $p < .001$, the 1.2-s exposure duration, $t(11) = 7.65$, $p < .001$, and the 4.2-s exposure duration, $t(11) = 4.32$, $p < .005$.

Free recall. As indicated earlier, subjects who judged letter frequency were subsequently asked to recall as many strings as possible. Our reason for doing this was to determine whether subjects' letter-frequency judgments were derived from activated memory representations for previously seen strings. For each letter we computed the proportion of subjects for whom the letter appeared in at least one correctly recalled string. As can be seen from the mean values in Table 4, we obtained the expected difference in probability of recall between high- and low-frequency letters; high-frequency letters appeared more often in correctly recalled strings than did low-frequency letters. This difference was expected because high-frequency letters occurred in more strings than low-frequency letters. What was important, however,

was that this difference was virtually identical when averaged over string frequency, for each of the three exposure durations.[2] An ANOVA, with the letters used in the experiment serving as the random variable, indicated that strings with high-frequency letters were recalled significantly more often than strings with low-frequency letters, $F(1, 14) = 104.57$, $p < .001$, $MS_e = .007$. The effect of exposure duration on string recall was significant, $F(2, 28) = 27.24$, $p < .001$, $MS_e = .017$, but neither the interaction between letter frequency and exposure duration, $F(2, 28) < 1.0$, $MS_e = .017$, nor the interaction between letter frequency and string frequency, $F(1, 14) = 2.13$, $p > .05$, $MS_e = .007$, was significant. Finally, the three-way interaction between exposure duration, letter frequency, and string frequency was significant, $F(2, 28) = 5.44$, $p < .02$, $MS_e = .017$. There was no obvious explanation for this interaction. It is worth noting, however, that the data pattern leading to the interaction did not match the pattern obtained for judgments of letter frequency.

The results obtained from this analysis of subjects' recall protocols stand in contrast with the significant interaction between letter frequency and exposure duration that was obtained when we assessed subjects' judgments of frequency of occurrence for individual letters. Subjects' frequency judgments differentiated among the high- and low-frequency letters for the 4.2-s and 1.2-s, but not for the 0.2-s exposure duration. This interaction was not obtained for the recall data. Subjects recalled more strings with high-frequency letters than strings with low-frequency letters, even at the briefest exposure duration. If subjects' frequency judgments for letters were based on whether or not they could recall a letter string containing the letter being judged, they would have been as accurate discriminating letter frequency for the 0.2-s exposure duration as they were for the 1.2-s and 4.2-s exposure durations. The case was particularly clear for letters from high-frequency strings. For 0.2-s exposures, subjects were clearly unable to judge the frequency of occurrence for these letters (see Table 2). However, it can be seen from the recall analysis, summarized in Table 4, that this combination of conditions (i.e., high-frequency strings, 0.2-s exposures) provided the greatest potential for subjects to base discriminative letter-frequency judgments on the activation of memory representations for previously seen strings. Yet, successful letter-frequency discrimination was not obtained.

---

[2] A separate analysis was performed for strings that were listed in subjects' recall but did not correspond with strings that were previously presented. High- and low-frequency letters appeared equally often in these incorrectly recalled letter strings.

237   H. HOCK, L. MALCUS, AND L. HASHER

An alternative version of the strategy discussed above is one in which frequency judgments are based on whether or not subjects can recall more than one string containing the letter being judged. Accordingly, the likelihood of a subject judging that a letter occurred 12 rather than 6 times would increase when the subject recalled more than one string containing the letter. As above, the potential utility of this strategy stems from high-frequency letters occurring in more different strings than low-frequency letters. Multiple recall was therefore more likely for the high- than the low-frequency letters.

In parallel with the preceding analysis, we computed the proportion of subjects for whom each letter appeared in at least two correctly recalled strings. In contrast with the preceding analysis, a letter appeared in more than one correctly recalled string very infrequently. The mean probabilities of multiple-letter recall, averaging across high- and low-frequency letters, were .02, .06, and .17 for the brief, intermediate, and long presentation durations, respectively. These probabilities were too low to permit the comparison of high- and low-frequency letters that was the basis for the preceding analysis. Instead, we deleted subjects' frequency judgments for letters that subsequently appeared in more than one of their correctly recalled strings. The results for judgments of letter frequency remained the same. Thus, there was no support for the hypothesis that subjects' frequency judgments for letters were based on the number of strings they could recall that contained the letter being judged.

Experiment 1A

The evidence obtained when the exposure duration was 0.2 s showed that subjects could accurately judge string frequency under conditions that prohibited accurate discrimination of letter frequency. The purpose of this experiment was to replicate that finding.

Method

As indicated above, the exposure duration was 0.2 s per string (the time between exposures remained at approximately 0.8 s). The design and procedure were identical to Experiment 1. A group of 32 subjects voluntarily participated in the experiment. All were students in an undergraduate psychology class at Florida Atlantic University. Half judged letter frequency and half judged string frequency.

Results

Letter-frequency judgments. The overall percentage of letter-frequency judgments that were correct was 52%, which was effectively at chance. The proportions of responses for which the letters were judged to have occurred 12 times are presented in Table 2. As in the brief exposure condition of Experiment 1, subjects' judgments did not differentiate between letters presented 12 times and letters presented 6 times (their proportion of correct frequency judgments was at chance). This was the case for letters from both high- and low-frequency strings. An ANOVA on the proportion of "12" responses in subjects' letter judgments indicated that the effect of letter frequency, $F(1, 15) < 1.0, MS_e = .042$, the effect of string frequency, $F(1, 15) < 1.0, MS_e = .090$, and the interaction between letter frequency and string frequency, $F(1, 15) < 1.0, MS_e = .055$, were not significant. Identical results

were obtained when the letters used in the experiment replaced subjects as the random variable in the ANOVA. The effects of letter frequency, string frequency, and the interaction between letter frequency and string frequency were not significant, $F(1, 14) < 1.0, MS_e = .024$ (for all of the above).

String-frequency judgments. The overall percentage of string-frequency judgments that were correct was 66%. The proportions of responses for which the strings were judged to have occurred six times are presented in Table 3. As in Experiment 1, subjects' responses differentiated between the strings presented six times and the strings presented three times. An ANOVA indicated that the effect of string frequency on the proportion of "6" responses in subjects' string judgments was significant, $F(1, 15) = 27.33, p < .001, MS_e = .058$. Although string frequency judgments were significantly higher for strings with low-frequency compared to high-frequency letters, $F(1, 15) = 5.06, p < .05, MS_e = .034$, the extent to which subjects' judgments discriminated between the two levels of string frequency was not affected by the frequency of the letters composing the strings. That is, the interaction between letter frequency and string frequency was not significant, $F(1, 15) < 1.0, MS_e = .023$. Similar results were obtained when items replaced subjects as the random variable in the ANOVA. The effect of string frequency was significant, $F(1, 14) = 11.42, p < .005, MS_e = .035$, and the interaction between string frequency and letter frequency was not significant, $F(1, 14) < 1.0, MS_e = .035$. In contrast to the analysis with subjects serving as the random variable, the main effect of letter frequency was not significant, $F(1, 14) = 1.17, p > .05, MS_e = .035$. The final analysis contrasted frequency judgments for high-frequency strings composed of low-frequency letters with low-frequency strings composed of high-frequency letters. Judgments were significantly higher for the high-frequency strings, $t(15) = 7.09, p < .001$.

Discussion

The results closely replicated those obtained in the comparable 0.2-s exposure duration of Experiment 1; subjects' judgments discriminated string frequency under conditions that prohibited the discrimination of letter frequency.[3] Subjects could not have

---

[3] A potential objection to this conclusion was that subjects may have had knowledge of letter frequency for briefly presented strings, but it was more difficult to discriminate 12 from 6 events, as required for individual letters, than it was to discriminate 6 from 3 events, as was required for strings. An additional experiment was conducted to evaluate this possibility. The stimuli in this experiment were individual letters, half of which were presented 12 times each and half presented 6 times each. The exposure duration was 0.2 s. The results of the experiment indicated that the difference in the proportion of "12" responses between the high-frequency letters (.65) and low-frequency letters (.32) was comparable to the difference in the proportion of "6" responses between the high- and low-frequency strings in the 0.2-s exposure conditions of Experiments 1 and 1A. We could therefore rule out the possibility that subjects failed to discriminate letter frequency for 0.2-s exposures in Experiments 1 and 1A because of the relative difficulty inherent in discriminating between 12 and 6 events. Although several explanations are possible, it seems most likely to us that differences in frequency judgments for strings and letters were due primarily to the relative ease of retaining information for one string relative to retaining information for four independent.

based their judgments of string frequency on the frequency of occurrence of their component letters because they apparently had no knowledge of the letter frequencies.

One aspect of the data that is difficult to explain is that strings tended to be judged higher in frequency when they were composed of low-frequency compared with high-frequency letters. This difference, which was not significant in Experiment 1, did not directly contribute to the discrimination of letter frequency because string frequency and letter frequency were varied orthogonally in both experiments. In addition, the direction of the difference was opposite what would be expected if the presence of high-frequency letters in a string biased subjects to judge the string as high in frequency. Because it was not reliable when subjects was the random variable in the analysis, it may be that the effect of letter frequency on string frequency judgments was due to uncontrolled differences in familiarity among some of the strings.

Experiment 2

The purpose of this experiment was to replicate and generalize the finding that letter-level frequency information can be abstracted from sequences of strings containing the letters. The replication extended the results of Experiment 1 in the following ways: (a) Subjects' task during the initial presentation of the strings involved making a lexical (word/nonword) decision for each item rather than remembering the information presented on the screen, (b) there were four rather than two frequency levels for the target letters, (c) instead of immediate testing, a 30-min delay was introduced between the initial presentation of the strings and the letter-frequency test, and (d) a partial-correlation procedure was used to show that subjects' ability to estimate letter frequency did not depend on their frequency estimates being based on the number of strings that they could recall that contained the letter being judged.

Method

Subjects. Sixteen undergraduate students in psychology classes at Florida Atlantic University participated in the experiment, for which they received class credit.

Stimuli. Sixteen consonants were selected as target letters for the experiment. For one stimulus set, eight of the target letters appeared only in common, four-letter words and the other eight appeared only in pronounceable four-letter nonwords. Each set of eight letters was further subdivided into four subsets of two letters, with each subset assigned to one of four frequency levels (4, 8, 16, and 32). The average frequencies of usage (i.e., the background frequencies) were similar for each pair of letters (based on the Mayzner & Tresselt, 1965, norms). For the second stimulus set, the eight target letters appearing in words and the eight target letters appearing in nonwords were switched. The frequency levels to which the letters were assigned were also changed.

Every string contained either one or two target letters. A target letter never appeared more than once in any string, and each string appeared only once in each stimulus set. One stimulus set had 100 words and 100

simultaneously occurring letters. Regardless of why frequency judgments failed to discriminate between the two levels of letter frequency (for letters embedded in strings) under conditions for which they successfully discriminated between the two levels of string frequency, we could conclude that subjects' frequency judgments for strings were not derived from the frequencies of their constituent letters.

nonwords. The other stimulus set had 40 more filler items (evenly divided between words and nonwords), which were strings composed entirely of nontarget letters. The strings belonging to each stimulus set were presented in two random orders (one order was the reverse of the other), resulting in four different stimulus sequences. Each subject was presented one of these sequences. Letter repetitions were separated in each stimulus sequence by a minimum of two intervening strings. Preceding each experimental sequence were 40 randomly ordered practice strings (half words, half nonwords). None of the practice strings was presented more than once, and none included any of the target letters.

Procedure. The stimuli were displayed on an Electrohome black and white television monitor that was controlled by a Data General Eclipse computer. Each string was presented inside a small rectangular box that always remained on the screen. The box intercepted a visual angle of 1.1° vertically and 2.8° horizontally. Each string intercepted a visual angle of 0.7° vertically and 2.4° horizontally. The exposure duration for each string was 1 s. The interstimulus interval was 1 s, except for the occasional trials for which the subject required more than 1 s to respond. Then, a 1-s delay was introduced between the subject's response and the presentation of the next stimulus. Subjects were provided with a response box and instructed to press the button marked "yes" if the string presented was a word. They were to press the button marked "no" if the string presented was not a word. Subjects were told to respond as quickly as possible, but to keep their errors to a minimum (incorrect responses were signalled by a brief flash of the rectangular box on the screen). They received no instructions suggesting that there would be any sort of memory test. When asked at the conclusion of the experiment, none of the subjects indicated that they expected to receive a memory test. Intervening between the lexical-decision task and the letter-frequency test was a 30-min interval during which subjects participated in a choice-reaction-time task involving judgments of visual direction.

During the letter-frequency test, subjects were presented with one of two random orders of all 26 alphabet letters (one order was the reverse of the other). Accompanying each letter was a number. This number was at the midpoint of the range of frequency values for all the target and nontarget letters appearing in the stimulus sequence. For two of the stimulus sequences, the range of letter frequencies was 1 to 90. For the longer two stimulus sequences the range was 1 to 130. Subjects were required to adjust the number appearing on the monitor upward or downward, using the same response buttons as in the preceding lexical-decision task, to reflect their estimate of how often each letter appeared over the entire sequence of strings. They were given all the time they needed for each response before the next letter was presented.

Results

Mean reaction times for subjects in the lexical-decision task were 632 ms for "yes" responses and 680 msec for "no" responses. The advantage in processing time for "yes" responses is typical of the lexical-decision paradigm.

Correlations between actual letter frequency and estimated letter frequency were computed for each subject. One correlation was computed for the eight target letters assigned to the words and another for the eight target letters assigned to the nonwords. The mean correlations were .41 for the words and .40 for the nonwords. The overall mean was significantly greater than zero, $t(15) = 5.78$, $p < .001$ (of the 32 correlation coefficients that were computed, 28 were positive).[4]

---

[4] Correlation coefficients were used as descriptive statistics in this experiment. The computation of mean correlation coefficients and statistical analyses of individual correlation coefficients were based on Fischer's *r* to *z* transformation.

239   H. HOCK. L. MALCUS. AND L. HASHER

As in Experiment 1, recall protocols were used to evaluate the likelihood that subjects' letter-frequency judgments were based on the retrieval of global-level memory representations of the strings containing the letters. We adopted the criterion that three of the four letters in the string had to be correct before the string was considered to be correctly recalled and then counted the frequency with which each target letter appeared in a subject's correct-recall protocols. These recall frequencies were correlated with both the actual and estimated letter frequencies, and a partial correlation coefficient was computed between the actual and estimated frequencies. Using this procedure. the relationship between actual and estimated frequency was determined. with the effects of recall frequency held constant (McNemar. 1962).

Partial correlation coefficients for each subject were computed separately for the eight target letters assigned to the word and the eight target letters assigned to the nonwords. The mean correlation coefficients. .46 for the words and .40 for the nonwords, were not statistically different, $t(15) < 1.0$. However, the overall mean, $r = .43$, was significantly greater than zero, $t(15) = 4.75$, $p < .001$ (of the 32 partial correlation coefficients that were computed, 27 were positive).

Discussion

The results of this experiment extended the previously reported evidence for the abstraction of letter-level frequency information. Subjects' estimates discriminated between the frequencies of letters as they appeared in both words and pronounceable nonwords during a lexical-decision task. The partial correlation procedure showed that subjects' success at letter-frequency estimation could not be attributed to the derivation of their estimates from the recall of the strings containing the letters. Furthermore. the introduction of a 30-min delay between the presentation of the strings and the letter-frequency test showed that the letter-frequency information was stored in long-term memory. This result paralleled Warren and Mitchell's (1980) evidence for the retention of string-level frequency information over a 20-min delay. Substantial losses in string-level frequency information have been reported for 1-week delays (Underwood. Zimmerman. & Freund. 1971). Whether or not extended delays would result in similar losses in abstracted letter-level frequency information remains to be determined.

General Discussion

Our basic evidence for the storage of element-level units was that high-frequency letters were judged as high in frequency significantly more often than low-frequency letters (for exposure durations of 1.2 and 4.2 s). We then considered the counterargument that frequency information was not associated with letter-level memory units and that subjects based their letter judgments on the activation of global-level memory units containing the letters. We noted first that the stimulus list was designed such that letter and string frequency were varied orthogonally, so that string frequency had no predictive value for judgments of letter frequency. Furthermore, high-frequency letters were judged as higher in frequency than low-frequency letters (for the 1.2- and 4.2-s durations) even when the high-frequency letters appeared in low-frequency strings and the low-frequency letters appeared

in high-frequency strings. Because differences in string frequency would have favored the activation of global-level units with low-frequency letters. this comparison provided evidence against the argument that element-level frequency judgments depended on the activation of global-level memory units. The final evidence for this conclusion came from an analysis of subjects' recall protocols. As expected from the structure of the stimulus list, subjects recalled more strings with high-frequency than low-frequency letters (there were more of the former to recall). However. this was the case even at the 0.2-s exposure duration. Despite this indication of greater availability of global-level memory units containing high-frequency letters compared with low-frequency letters, subjects still failed to discriminate between high- and low-frequency letters following 0.2-s exposures of the strings. It could be concluded. at least for the longer exposure durations of Experiment 1, that subjects abstracted letter-level frequency information that characterized the componential structure of the stimulus list. Further evidence indicating that this information was stored in long-term memory was reported in Experiment 2.

The letter judgment data provided some indication that Tversky and Kahneman's (1973) "availability heuristic" was operative in our experiments. That is, there was some bias for subjects to judge letters as relatively high in frequency simply because they occurred in more available high-frequency strings (the main effect of string frequency on letter judgements was significant in Experiment 1 but fell short of significance in Experiment 1A). However, the orthogonal manipulation of letter and string frequency prevented this bias from being the basis for subjects' discrimination between high- and low-frequency letters.

Experiments involving judgments of string frequency generally do not control for the frequency of occurrence of the letters composing the strings. The above evidence for letter-frequency discrimination suggests the possibility that frequency judgments for strings might be based. not on their frequency of occurrence. but on the frequency of occurrence of their constituent letters. Although the use of such a strategy is possible. the stimulus list was designed so that string frequency could not be predicted from letter frequency (they were varied orthogonally). Furthermore. high-frequency strings were judged as higher in frequency than low-frequency strings even when the former were composed of low-frequency letters and the latter of high-frequency letters. The opposite result would have been obtained if judgments of string frequency were based on the frequencies of the strings' constituent letters. Finally, string-frequency judgments discriminated between high- and low-frequency strings under experimental conditions (the 0.2-s exposure duration) for which letter-frequency judgments did not discriminate between high- and low-frequency letters. Under these conditions. it was impossible for string frequency to be judged on the basis of stored information involving letter frequency because the latter information was not available (as measured by judgments of letter frequency).

To summarize. our experimental results provided evidence for the independent formation of both global-level (strings) and element-level (letters) memory units. Although this conclusion refers to the sort of information subjects remember, the results are potentially informative concerning the way in which this information is retrieved. That is, the tendency for string-frequency information to bias letter-frequency judgments suggests that both types of information may be retrieved together, even when the

task calls only for the judgment of letter frequency. (The case for joint retrieval when subjects judge string frequency remains uncertain because these judgments appeared to be influenced only by string-level information.) The joint retrieval of string-level and letter-level memory units might suggest, but does not demand, that they are stored together. One possibility is that representations involving global-level (string) memory units include individual element-level units (the spelling for the string), with each of the elements tagged with regard to its frequency of occurrence in the list. Whether one can distinguish between this and other representational formats (e.g., separate storage of global-level and element-level units) is questionable. This indeterminacy is analogous to the indeterminacy regarding the representational format for item-specific, exemplar information (global-level units in our study) and category-level, summary information (element-level units in our study) in category-acquisition tasks (Medin, Dewey, & Murphy, 1983). In the absence of experimental paradigms that definitively distinguish among alternative representational formats, our further research involving the formation of memory units has replicated the results reported in this article and focused on the task characteristics that influence the retention of informational units of different size. For example, we are finding that the phonemic processing of strings of letters facilitates the formation of memory units for the letters (Malcus, Hock, Cavedo, & Smith, 1985).

Evidence that subjects can accurately judge the frequency of occurrence of component elements of larger order units is not exclusive to this article. Jacoby (1972) showed that subjects can judge the frequency of occurrence of words embedded in grammatical sentences. Investigators testing feature-frequency models of concept formation have shown that subjects can judge frequency of occurrence for the component parts of schematic faces (Kellogg, 1981). In work in progress, we are obtaining evidence that subjects presented with a sequence of dot patterns can accurately judge how often individual locations have been occupied by a dot. In other work in progress, we are finding that subjects can discriminate differences in the frequency with which various spatial relations (e.g., inside, below) occur across a series of scenes involving different objects. However, obtaining evidence for successful frequency discrimination does not definitively indicate that subjects have stored frequency information regarding elements and relations between elements. It may be necessary to show, as in the present article, that subjects' frequency judgments for the componential characteristics of a series of stimuli are not based on the activation of global-level memory units corresponding to the particular stimuli presented in the series.

The frequency-judgment procedure seems to be particularly useful for assessing memory units, especially when the vocabulary of elementary units is limited, as is the case for alphabet letters. For example, when frequency judgments are used to assess memory for the constituent letters in strings, the full set of alphabet letters can be presented during the acquisition phase of the experiment. In contrast, recognition and recall procedures would have to exclude a significant proportion of the alphabet from the acquisition phase because they must be used to detect false recognition responses or recall intrusions during the subsequent memory test. Furthermore, the frequency-judgment

procedure eliminates the need for production (as in recall paradigms) and also eliminates the problem of performance depending on the perceptual similarity of previously seen and new distractor stimuli (as in recognition paradigms).

References

Attneave, F. (1953). Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology, 46.* 81–86.

Burnett, S. A., & Stevenson, J. H. (1979). Instructional effects on frequency judgments for sentence meanings. *American Journal of Psychology, 92.* 711–721.

Carroll, J. B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior, 10.* 722–729.

Gude, C., & Zechmeister, E. B. (1975). Frequency judgments for the 'gist' of sentences. *American Journal of Psychology, 88.* 385–396.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108.* 356–388.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88.* 297–306.

Hochberg, J. (1981). Levels of organization. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization.* Hillsdale, NJ: Erlbaum.

Howell, W. C. (1973). Storage of events and event frequency: A comparison of two paradigms in memory. *Journal of Experimental Psychology, 98.* 260–263.

Jacoby, L. L. (1972). Context effects on frequency judgments of words and sentences. *Journal of Experimental Psychology, 94.* 255–260.

Kellogg, R. T. (1981). Feature frequency in concept learning: What is counted? *Memory & Cognition, 9.* 157–163.

Malcus, L., Hock, H. S., Cavedo, L. C., & Smith, L. B. (1985, March). *Orthographic processing of strings affects frequency discrimination for component letters.* Paper presented at the meeting of the Eastern Psychological Association, Boston, MA.

Mayzner, M. S., & Tresselt, M. E. (1965). Tables of single letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1*(No. 2), 13–78.

McNemar, Q. (1962). *Psychological statistics.* New York: Wiley.

Medin, D. L., Dewey, G. I., & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9.* 607–625.

Neisser, U. (1967). *Cognitive psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Rowe, E. J. (1973a). Context effects in judgment of frequency. *Bulletin of the Psychonomic Society, 2.* 231–232.

Rowe, E. J. (1973b). Frequency judgments and recognition of homonyms. *Journal of Verbal Learning and Verbal Behavior, 12.* 440–447.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5.* 207–232.

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76.* 559–573.

Underwood, B. J., Zimmerman, J., & Freund, J. S. (1971). Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology, 87.* 149–162.

Warren, L. R., & Mitchell, S. A. (1980). Age differences in judging the frequency of events. *Developmental Psychology, 16.* 116–120.

SECTION 2B

Frequency Estimation for Words Appearing in Sentences

Howard S. Hock and Mary LaLomia

An issue of particular importance in recent research concerned with the encoding of frequency of occurrence information is whether or not the encoding process is automatic (Hasher and Zacks, 1979; 1984). Although it is clear that frequency information can be encoded without the intention to do so (Hasher, Zacks, Rose, & Sanft, 1986), experiments demonstrating that subjects' orienting task in viewing a list of words affects the accuracy of frequency judgments suggest that the encoding process may not be not automatic. For example, Greene (1984) has reported an advantage in frequency judgment for subjects receiving an intentional-learning orienting task compared with subjects receiving an incidental-learning orienting task (without any instructions suggesting that subjects remember the presented items). In another study, Naveh-Benjamin and Jonides (1986) found, for subjects with an intentional-learning orienting task, that frequency judgments were more accurate when a secondary orienting task was relatively difficult compared with when it was relatively easy. They also found that frequency judgments were more accurate following a semantic-association orienting task than they were following an acoustic-association orienting task.

How can frequency coding be both unintentional and task dependent? The explanation we propose is that subjects' attention to an item during the orienting task can affect the likelihood of previously seen versions of the item being activated in memory. Once attention is paid to an item, the reactivations that facilitate frequency discrimination proceed without trying to remember how frequently various events occurred.

The "attention hypothesis" was tested by manipulating the extent to which subjects attended to the items whose frequency they would subsequently be judging. Rather than having subjects judge the frequency with which various words appearing in a list of words, we asked subjects to judge the frequency of occurrence of target nouns that appeared in a series of sentences. One group of subjects judged whether or not the sentences were meaningful. A second group of subjects judged whether the sentences were written in the present or past tense. Because the judgment of meaning demanded attention to the nouns in each sentence, but the judgment of tense did not, we predicted that the estimation of target-noun occurrence frequency would be more accurate for the meaning-judgement group than the tense-judgment group.

The present study was concerned with quantitative differences in code activation resulting from differences in attention during the initial presentation of the items. However, previous experiments have suggested that orienting tasks of the type we used could result in qualitative rather than quantitative differences in code formation. Bransford, Franks, Morris, and Stein (1979) have argued that the apparent effectiveness of different orienting tasks depends on the types of codes that are formed and the requirements of the post-acquisition memory test. For example, phonological orienting

tasks result in better post-acquisition recognition accuracy than semantic orienting tasks when the previously seen and distractor items are phonologically similar, but vice versa when they are semantically similar. However, as indicated earlier, an advantage of frequency-estimation tests of memory is that they do not require the presentation of distractor items. Hence, frequency estimation is neutral with regard to qualitative differences in the codes subjects generated during the initial presentation of the items, and reflect only quantitative differences in code formation. These quantitative differences are presumably due to differential attentional requirements of different orienting tasks.

A further consideration in evaluating the effect of the orienting task on word-frequency estimation concerns the size of the memory units on which frequency estimates are based. In the Hock, Malcus, and Hasher (1986) study, we were concerned with whether letter frequency estimates were based on the retrieval of letter-level or string-level memory units. That is, in addition to the retrieval of letter-level memory codes, subjects could base their estimate of a letter's frequency on the number of global-level memory units they could recall that included the letter being estimated. Hock, et al. (1986) evaluated this possibility by obtaining free recall data following the frequency estimation phase of their experiment, determining the number of times each letter appeared in correctly recalled strings, and computing partial-correlation coefficients with the effects of recall-frequency "held constant." A similar procedure was used in the present study to determine whether judgments of word frequency were influenced by the recall of sentence-level memory units.

Experiment 1

Method

Subjects. Sixty four undergraduate students in psychology classes at the University of South Florida participated in this experiment, for which they received class credit.

Stimuli. Sixteen target nouns were chosen for the experiment. Their frequency of occurrence in written English was high; the mean frequency value was 268 per million words. For each target noun, 16 different sentences were constructed, eight meaningful and eight meaningless. Half of the meaningful sentences were written in the present tense, half were written in the past tense. The same was true for the meaningless sentences. The meaningless sentences were constructed from the meaningful ones by replacing a noun or verb such that the meaningless sentences remained grammatically correct. For example, the meaningful sentence THE TAXI WENT SPEEDING BY was rendered meaningless by replacing the word SPEEDING with the word ITCHING. The target noun in this example was TAXI. The positions of the target noun and the tense-determining verb were balanced across sentences.

Although a total of 256 sentences were constructed (16 target nouns x 16 sentences), each subject was presented only 60 of the sentences. Of these 60 sentences, 30 were meaningful (15 were in the present tense, 15 were in the past tense), and 30 were meaningless (15 were in the present tense, 15 were in the past tense). Four of the 16 target nouns each appeared in eight different sentences, four appeared

in four different sentences each, four appeared in two different sentences each, and four appeared in one sentence each. At each frequency level, half the target nouns appeared in meaningful sentences and half appeared in meaningless sentences. If a particular target noun appeared in a meaningful sentence, its other appearances in a list could only be in other meaningful sentences (the same was true for target nouns appearing in meaningless sentences). Thirty two different lists were constructed so that each target noun appeared equally often at each frequency level (1, 2, 4, and 8), in each sentence type (meaningful vs. meaningless), and with each tense-type (present vs. past). Over the entire experiment, each of the 256 sentences appeared equally often.

Design. The experiment was conducted in two phases. There were two orienting tasks during Phase 1, Meaning-Judgment and Tense-Judgment. The 32 subjects assigned to the Meaning-Judgment condition were required to determine whether each sentence was meaningful or meaningless. The 32 subjects assigned to the Tense-Judgment condition were required to determine whether the sentences were written in the present or past tense. Subjects were given no indication that there would be any sort of memory test. Phase 2 constituted a frequency estimation test, with each subject being presented one of four different random orders of the 16 target nouns. Post-experimental interviews confirmed that the memory test of Phase 2 was unexpected.

Procedure. The stimuli were displayed on an Apple 2e microcomputer. Each sentence was presented inside a rectangular frame that always remained on the screen. The frame intercepted a visual angle of 2 deg vertically and 20 deg horizontally. The exposure duration for each string was three sec. The interstimulus interval was also one sec, except for the occasional trials on which subjects required more than one sec to respond. Then, a one sec delay was introduced between the subject's response and the presentation of the next stimulus. Subjects in the Meaning-Judgment condition were instructed to press one button on the computer keyboard for sentences that could corresponded to a real event, and another button otherwise. Subjects in the Tense-Judgment condition were instructed to press one button for sentences that were in the present tense, and another button for sentences that were in the past tense. Incorrect responses were signalled by a brief flash of the rectangular box on the screen. During Phase 2, subjects were asked to estimate how often each target word appeared in the preceding list. The to-be-judged words appeared one at a time on the screen.

Results

Mean reaction times and error rates for Phase 1 are presented in Table 1. An examination of the table indicates that reaction times were faster in the Tense-Judgment than the Meaning-Judgment condition, but the difference was probably due to differential speed-accuracy criteria in the two conditions. The finding of primary interest was that reaction times in the Tense-Judgment condition were faster for the meaningful than the meaningless sentences. This indicated that subjects in this condition were processing the meaning of the sentences even though the task required processing only verb tense.

Table 1
Mean Reaction Times (in msec) and Error Rates (pct) for the Phase 1 Orienting Tasks of Experiment 1.

Phase 1 Task	REACTION TIMES		ERROR RATES	
	Meaningful Sentences	Meaningless Sentences	Meaningful Sentences	Meaningless Sentences
Meaning-Judgment	723	813	3.8	11.3
Tense-Judgment	666	771	12.2	14.2

The results for the frequency estimation test of Phase 2 are presented in Table 2. As predicted, the Meaning-Judgment orienting task resulted in more accurate estimation of word frequency than the Tense-Judgment orienting task. This supported the hypothesis that the encoding of information relevant to word-frequency judgments was facilitated by tasks requiring attention to the to-be-judged items; since determining the tense of verbs did not require attention to the target nouns in the sentences, frequency judgments were relatively inaccurate in the Tense-Judgment condition.

Table 2
Mean correlation coefficients between estimated and actual word frequency for Experiment 1.

Phase 1 Task	Meaningful Sentences	Meaningless Sentences
Meaning-Judgment	0.62	0.52
Tense-Judgment	0.42	0.42

Experiment 2

This experiment was similar to the Meaning-Judgment condition of Experiment 1, with the exception that the materials were different and the experiment included a third phase in which subjects were given 10 min to recall as many of the sentences as possible. The purpose of the experiment was to determine whether judgments of word frequency for words appearing in a series of sentences depended on the retrieval of word-level or sentence-level memory units. Thirty two undergraduate students at Florida Atlantic University participated in this experiment as one way of earning credit toward an undergraduate psychology course.

Table 3
Mean correlation coefficients between estimated and
actual word frequency, and between estimated and actual
word frequency with recall-frequency partialled-out.

	Meaningful Sentences	Meaningless Sentences
Actual with Estimated Frequency	0.48	0.55
Actual with Estimated Frequency (with frequency in correctly recalled sentences partialled-out)	0.47	0.53
Actual with Estimated Frequency (with frequency in all recalled sentences partialled out)	0.42	0.50

The results of the experiment are summarized in Table 3.
Partial-corelation coefficients were computed as follows. In the
initial analysis, the sentences subjects recalled during Phase 3 were
scored as correct if the critical information in the sentence was
correctly recalled; erroneous recall of articles or the tense of the
verbs was ignored. Using this criterion, correct recall was quite
low: 5.6% for the meaningful sentences, 1.7% for the meaningless
sentences. We then counted the frequency with which each target word
appeared in correctly recalled sentences and computed
partial-correlation coefficients. As indicated in Table 3, the
partialling procedure had no appreciable effect on the correlations
between actual and estimated frequency. In an additional analysis, we
eliminated the criterion that the sentences had to be correctly
recalled, and simply counted the frequency with which each target word
appeared in subjects' recall. The partial-correlations computed on
the basis of recall-frequency were only slightly reduced. We
concluded, therefore, that subjects' estimates of word frequency for
words appearing in sentences are based on the retrieval of word-level
as opposed to sentence-level memory units.

References

Bransford, J.D., Franks, J.J., Morris, C.D., & Stein, B.S. (1979).
    Some constraints on learning and memory research. In L.S. Cermak
    and F.I.M. Craik (Eds.), Levels of processing and human memory.
    Hillsdale, N.J.: Erlbaum.

Greene, R.L. (1984). Incidental learning of event frequency. Memory &
    Cognition, 12, 90-95.

Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes
    in memory. Journal of Experimental Psychology: General, 108,
    356-388.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information: The case for frequency of occurrence. American Psychologist. 39, 1372-1388.

Hasher, L., Zacks, R.T., Rose, K.C., & Sanft, H. (1987). Truly incidental encoding of frequency information. American Journal of Psychology. in press.

Hock, H.S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. Journal of Experimental Psychology: Learning. Memory. & Cognition. 12, 232-240.

Naveh-Benjamin, M. & Jonides, J. (1986). On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. Journal of Experimental Psychology: Learning. Memory. & Cognition. 12, 378-386.

Section 2C

The Perception of Pattern: Coding the
Position of Component Elements

Howard Hock, Laurel Smith, Leonie Escoffery and Alexandra Bates

Abstract

When presented a series of patterns inside a frame, subjects unintentionally retained information concerning the location of the elements composing the patterns. Despite their certainty that they lacked the ability to do so, subjects could use this information to estimate the frequency with which the elements occurred at various locations within the frame. The precision with which subjects encoded the positions of individual elements was increased by orienting tasks which enhanced the formation of pattern-level memory units. Nonetheless, the accurate estimation of element-location frequency depended on the retrieval of element-level rather than pattern-level memory units. Orienting tasks that emphasized the formation of pattern-level memory units reduced estimation accuracy by suppressing the formation of element-level memory units. It was therefore concluded that the retention of superficial details, like the positions of the constituent elements of a pattern, need not be retained at the pattern-level in the form of a template-like pictorial copies. Element-level memory units, and their associated position codes, can be abstracted from a series of patterns and represented in memory in the same way as the semantically important attributes of the patterns.

Some time ago, Clement and Weiman (1970) reported the results of a study which suggested that it was extremely difficult, if not impossible, for subjects to ignore the configuration of a pattern composed of five filled circles and identify it strictly on the basis of whether or not a single spatial location was occupied by one of the circular elements. Although many years have passed since this study was published, there has, to our knowledge, been no further research concerned with the relationship between the encoding of pattern-level configural information and element-level location information.

Clement and Weiman (1970) required subjects to sort cards containing one of two patterns into two piles. When they were instructed to attend to the whole pattern, sorting times were affected by the size of the patterns' equivalence set; patterns with small equivalence sets (the configurally simpler patterns) were easier to discriminate than patterns with large equivalence sets (the configurally more complex patterns). The critical feature in Clement and Weiman's (1970) experiment was that they selected their pattern-pairs such that subjects could discriminate between the two patterns in each deck by attending to one location and determining whether or not an element (i.e., a filled circle) was present at that location. When subjects were told to look for such a location during the sorting task, the effect of equivalence set (configuration) was reduced, but not eliminated. When subjects were actually shown a location for which the presence or absence of an element discriminated between the two patterns, the configural effect was further reduced, but once again, was not eliminated.

Clement and Weiman's (1963) results provide evidence of a bias for patterns to be processed at a configural, or pattern-level, rather than element-level (Navon, 1977). There are a number of ways this bias could have influenced performance. One possibility is that subjects occasionally processed the stimuli at the pattern-level, despite intending to follow instructions to sort the patterns on the basis of element-location. Since some responses would be based on pattern-level processing, the configural effect would be reduced, but not eliminated by element-level instructions. Although the preceding explanation implies that pattern-level and element-level processing compete with each other, an alternative explanation implies cooperativity. That is, pattern-level information may always be processed, despite subjects' intentions to do otherwise. The residual configural effect obtained under element-level instructions could then be attributed to patterns with simple configurations being identified fast enough to facilitate a pooled decision that takes account of the "status" of both pattern-level and element-level processing channels (Miller, 1981).

The research reported in this study was aimed at furthering our understanding of the relationship between the encoding of pattern-level and element-level location information. In contrast with Clement and Weiman's (1970) research, it emphasized the retention rather than the processing of element-level information. As in studies concerned with the processing of pattern-level vs. element-level information (Navon, 1977; Pomerantz & Sager, 1975), the ideal of global precedence in the retention of information has intuitive appeal. In the case of printed words, meaning is not conveyed by individual letters, but by the

orthographic relations among the letters constituting the words. In the case of patterns, meaning is not conveyed by individual elements, but by the spatial relations among the elements composing the patterns. From a point of view that stresses encoding abstract, higher-order units that represent semantically important characteristics of a stimulus (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), there would be no reason to expect element-location information to be retained; the configural aspects of the pattern should take precedence with respect to the retention of pattern information.

The above arguments notwithstanding, a number of experiments have demonstrated that subjects unintentionally retain what appear to be meaningless, superficial aspects of various stimuli. For example, Jacoby and Brooks (1984) have demonstrated that the speed with which subjects can identify a picture is enhanced when the same picture had been seen on a previous occasion; the repetition of superficial details that were irrelevant to the identification of the picture facilitated its identification the second time it was presented. Similarly, Pollatsek, Rayner, and Collins (1984) have shown that identification time for a foveally viewed picture is enhanced when the exact same picture had been viewed parafoveally a moment before. Many different experiments have provided evidence for the retention of information concerning the case in which letter strings were printed (e.g., Hock, Throckmorton, Rosenthal, & Webb, 1981). Kolers (1976) has demonstrated that the speed with which subjects re-read a series of passages was enhanced by exact repetition of the typographical characteristics of the passages, even though more than a year had passed since their initial reading.

Given this evidence for the unintentional encoding of superficial detail, as well as Clement and Weiman's (1970) evidence of a bias to process patterns at a configural level, the initial objective of the research reported in this paper was to determine whether subjects would unintentionally encode superficial characteristics of the patterns: the positions of their constituent elements. The experimental procedure we used to make this determination was based on a paradigm developed by Hock, Malcus, and Hasher (1986) to study the encoding of frequency of occurrence information for elements (letters) that were constituents of higher-order perceptual units (letter strings). Hock, et al. (1986) presented subjects a list of words and pronounceable nonwords, and showed that they could subsequently judge the frequency with which individual letters had appeared in the list. That is, even though meaning was most efficiently represented at the string-level, memory units resulting from the abstraction of letter-level information allowed subjects to successfully estimate the frequency with which individual letters appeared in the list. The experimental paradigm, when applied to pattern perception, involved presenting a series of patterns inside a surrounding frame, with an orienting task that provided no indication of the memory test to follow. Subjects were then asked to estimate the frequency with which the elements composing the patterns occurred in various locations within the frame.

One of the principal issues that arises in this paradigm concerns the manner in which subjects generate their frequency of occurrence

estimates (Hock, et al., 1986). One possibility is that their frequency estimates for each location are based on the retrieval of pattern-independent, element-level memory units. Another possibility is that subjects derive their frequency estimates for each location by determining how many pattern-level memory units they could retrieve with elements at the location being tested (i.e., via Tversky and Kahneman's [1973] availability heuristic). That is, a location would receive a relatively high frequency of occurrence estimate if the subject could retrieve a relatively large number of patterns with elements at that location.

A method for determining whether frequency estimates are based on the retrieval of element-level or pattern-level memory units has been introduced by Hock, et al. (1986). The method involves the following: 1) after the estimation phase of the experiment, subjects are asked to recall as many of the patterns as possible, 2) we then count the frequency with which each location within the frame is occupied by an element in a subject's correct-recall protocols, and 3) we compute a partial-correlation coefficient for each subject in which the recall-frequency for each location is "partialled out" of the correlation between the actual and estimated frequency. If a significant partial-correlation coefficient remains (with the effect of pattern-recall "partialled out"), it could be concluded that subjects' estimates were influenced by element-level memory units with associated codes specifying the position of the elements relative to the frame.

In Experiment 1, we tested for the unintentional encoding of superficial element-position information by having subjects estimate the frequency with which elements occurred at various locations within the frame. We determined whether the position information used for these judgments were stored in conjunction with element-level or pattern-level memory units, and developed a technique for measuring the accuracy of element-position coding. We went on to investigate the effects of intentionally trying to remember element-location frequency (Experiment 2), and whether the formation of pattern-level memory units is compatible with, or competes with the formation of pattern-independent, element-level memory units (Experiments 3 and 4).

Experiment 1

The purpose of this experiment was to provide evidence for the unintentional encoding of location for the constituent elements of a series of patterns composed of circular elements. During the initial presentation of the patterns, one group of subjects was instructed to determine, for each pattern, whether or not three of the circles were aligned vertically or horizontally. A second group of subjects was instructed to determine, for each pattern, whether or not three of the circles were aligned diagonally. These tasks were selected because they were expected to produce differences in performance consonant with the well-established advantage in the processing of vertical/horizontal compared with oblique orientations (i.e., the "oblique effect"; Appelle, 1972). We could then determine whether the oblique effect would carry over into the encoding of element-location information.

Method

Subjects. Sixty four undergraduate students at Florida Atlantic University participated in this experiment as one way of receiving course credit in an undergraduate psychology class.

Stimuli. The patterns were each composed of five empty circles whose locations were defined in terms of an imaginary 3x3 matrix of possible element locations. With the constraint that every row and column of the imaginary 3x3 matrix had to contain at least one circle per pattern, 17 different patterns were generated (Garner & Clement, 1963). The 3x3 matrix was located in one of four positions within a frame large enough to accomodate 16 possible element locations (4x4). That is, the 3x3 matrix was placed in one of the four quadrants of the larger 4x4 matrix. Each of the seventeen 3x3 patterns was presented twice, each time in the same quadrant within the 4x4 matrix.

We were unable to avoid compositional constraints which resulted in circles being located in the middle locations of the 4x4 matrix with the greatest frequency and in the corner locations of the 4x4 matrix with the least frequency. However, the orientation of each of the 17 Garner and Clement (1963) patterns and their location within the 4x4 matrix were selected to independently maximize the range of frequencies for the corner, side, and middle locations of the 4x4 matrix. The frequency with which each of the 16 locations in the 4x4 matrix was occupied by a circle is presented in Figure 1. Frequencies varied from 4 to 12 for the four corner locations, from 2 to 20 for the eight side locations, and from 6 to 28 for the four middle locations.

Insert Figure 1 about here

The set of 34 patterns described above were presented in one of four orientations: all in the arbitrarily defined 0 deg orientation (which determined the frequency distribution of Figure 1), all rotated 90 deg, all rotated 180 deg, or all rotated 270 deg. Separate groups of subjects viewed the patterns in one of the four orientations in order to balance out possible effects of top-down and/or left-right scanning biases on subjects' estimates of how often locations within the frame were occupied by circles.

The presentation of stimuli and the recording of responses were controlled by a Data General Eclipse computer. The stimuli, white patterns on a black background, were presented on an Electrohome monitor. A 9.0 cm x 9.0 cm white frame was always present in the center of the screen (the lines defining the frame were 0.1 cm thick). There were no grid lines presented inside the frame. The patterns, each composed of five unfilled circles with a diameter of 1.0 cm, were presented inside the frame.[1] When viewed from a distance of 42.5 cm, a visual angle of 12.1 deg was intercepted by the frame and a visual angle of 1.3 deg was intercepted by each circle.

Procedure. The experimental procedure consisted of three phases. During Phase I, subjects were presented with the 34 patterns in random order and were given one of two detection tasks. One group of 32 subjects was required to press one button if there were three circles within a pattern that were vertically or horizontally aligned (16 of the 34 patterns met this requirement) and the other button if there was no vertical or horizontal alignment among the circles. A second group of 32 subjects was required to press one button if there were three circles within a pattern that were aligned in either a left or right diagonal (12 of the 34 patterns met this requirement) and the other button if there was no diagonal alignment.

A small dot was placed in the center of three of the five circles composing each pattern in order to facilitate the detection of aligned circles (each dot had a diameter of one pixel). The three circles containing a dot remained the same for the two presentations of each pattern. In the Vertical/Horizontal condition, these dots were placed in vertically or horizontally aligned circles. For patterns without a vertical or horizontal alignment, the three dots were randomly assigned to three different circles in each pattern. In the Diagonal condition, the three dots were placed in three diagonally aligned circles. For patterns without a diagonal alignment, the three dots were randomly assigned to three different circles in each pattern.

The patterns were presented for one sec each. The intertrial interval was also one sec, except for the occasional trials on which subjects required more than one sec to respond. Then, a one sec delay was introduced between the subject's response and the presentation of the next stimulus. Subjects were instructed to respond as quickly and accurately as possible; they were not told to expect any sort of memory test. Response times (from the onset of each pattern) and errors were recorded during this phase of the experiment.

Estimates of element-location frequency were obtained during Phase 2 of the experiment. Sixteen different stimuli were presented during the estimation phase, each being a single circle inside the frame. The circle for each stimulus was located at one of the 16 possible locations in the 4x4 matrix (grid lines were not presented). The order of presentation for the 16 estimation stimuli was randomized. On each estimation trial, the number 15 was presented 7.0 cm below the center of the frame. This number represented the midpoint of the range of element-location frequencies (2 to 28). Subjects were told that the frequency with which a circle was presented at each of the indicated locations varied from 1 to 30, and were instructed, by pressing buttons, to change the number 15 on the screen to a number which reflected their estimate of element-location frequency over the full set of patterns they had just seen (out-of-range estimates were not accepted by the computer).

In Phase 3 of the experiment, subjects were provided with a sheet of paper with 20 empty 4x4 matrices (including grid lines) and given 10 min to recall as many of the patterns they had seen during Phase 1 as possible (they were required to provide five elements for each attempted reproduction). The data from this phase of the experiment were used to

determine whether subjects' frequency estimates for each location were derived from patterns they could recall that contained circles at the locations being tested.

Results

Performance in orienting tasks. Mean response times and error rates for Phase 1 are presented in Table 1. The faster responses obtained for Correct Detections compared with Correct Rejections were typical of the reaction time literature. Response times were significantly slower in the Diagonal than the Vertical/Horizontal condition, [$F(1,62) = 4.85$, $p <$ .05, MSe = 32,410]. This result was typical of the "oblique effect" (Appelle, 1972). Error rates were very similar in the two detection conditions.

Insert Tables 1 and 2 about here

Accuracy of frequency estimation. That subjects were capable of estimating element-location frequency was indicated by the correlations between the actual and estimated frequencies for each location in the 4x4 matrix, which were computed individually for each subject. Although these correlations were, on average, relatively small, it can be seen from Table 2 that they were positive for most of the subjects in the Vertical/Horizontal and Diagonal conditions. The means of the actual/ estimate correlations (see Table 2) were significantly greater than zero [$r = 0.34$, $t(31) = 5.14$, $p < .001$; $r = 0.33$, $t(31) = 5.72$, $p < .001$], for the Vertical/Horizontal and Diagonal conditions, respectively. The correlations obtained for the two conditions were not significantly different from each other, [$t(62) < 1.0$.].[2]

The derivational strategy. As indicated in the introduction, there were potentially two ways subjects could have estimated element-location frequency: 1) they could have retained pattern-independent memory units corresponding to individual circles and estimated frequency on the basis of these multiple element-level memory units (Hintzman & Block, 1971), and 2) when required to estimate element-frequency at a particular location, they could have recalled as many patterns as possible with a circle at that location and derived their estimate from the number of recalled patterns. The latter derivational strategy, a version of Tversky and Kahneman's (1973) availability heuristic, could be successful because a location becomes high in element-occurrence frequency as a result of being occupied by a circle for many different patterns. The possibility that subjects' frequency estimates were derived from recalled patterns was evaluated during Phase 3 by having subjects recall the patterns that had been presented during Phase 1.

Our analysis of the derivational strategy was based on only correctly recalled patterns, regardless of whether they were recalled in the correct quadrant of the 4x4 matrix. We counted the frequency with which each location was occupied by a circle in a subject's correct-recall protocols, and then computed a partial-correlation coefficient for each subject in which the recall-frequency for each location was "partialled out" of the correlation between the actual and estimated

frequency for each location. The mean partial-correlation coefficients were somewhat reduced from the correlation coefficients observed without partialling out recall-frequency, but correlations remained positive for most of the subjects (see Table 2). The means of the actual/estimate correlations (with recall-frequency "partialled-out"), which are also presented in Table 2, were significantly greater than zero [r = 0.26, t(31) = 4.03, p < .001; r = 0.28, t(31) = 4.81, p < .001], for the Vertical/Horizontal and Diagonal conditions, respectively. The partial correlations obtained for the two conditions were not significantly different from each other, [t(62) < 1.0]. Although there may have been some tendency for subjects to have derived their frequency of occurrence estimates for each location from the number of patterns they could recall with a circle at that location, a significant actual/estimate correlation remained after the contributions of the derivational strategy were "partialled-out."

Insert Table 3 about here

Recall accuracy. As indicated in Table 3, we found that recall was more accurate in the Diagonal condition (12.3% of the patterns were correctly recalled) than in the Vertical/Horizontal condition (8.1% of the patterns were correctly recalled). This difference was significant [t(62) = 2.00, p < .05]. Approximately half the correctly recalled patterns were recalled in their correct quadrant in the 4x4 matrix. Since the chance rate was .25, it could be concluded that subjects retained some information concerning the global locations of the correctly recalled patterns.[3]

Compositional constraints. The next analysis was concerned with whether subjects based their frequency of occurrence estimates for each location on their awareness of compositional constraints regarding where circles were most likely to be located (i.e., they were constrained to occur in the middle of the frame more often than in the corners of the frame). Subjects who were aware of this constraint could have done an adequate job of estimating element-location frequency without remembering anything about the information presented during Phase 1 of the experiment. However, subjects' awareness that the corner locations were constrained to be lower in frequency than the middle locations would not allow them to differentiate between the frequencies at each of the four corner locations or the frequencies at each of the four middle locations.

The first step in evaluating whether subjects' estimates were influenced by their awareness of compositional constraints was to determine the mean frequency estimate at each of the 16 locations in the 4x4 matrix. This was done independently for the 32 subjects in the Vertical/Horizontal condition and the 32 subjects in the Diagonal condition. We then computed the correlation between actual and mean-estimated frequency independently for the four corner locations, the eight side locations, and the four middle locations. The three correlations in the Vertical/Horizontal condition were all positive: 0.95, 0.72, and 0.21 for the corner, side, and middle locations, respectively. Since the number of degrees of freedom associated with

each correlation was too small to test their statistical significance, our determination of their reliability rested on the repetition of these positive correlations in the Diagonal condition. For the latter, the correlations were again positive: 0.99, 0.77, and 0.54 for the corner, side, and middle locations, respectively. We could conclude, therefore, that the observed correlations between actual and estimated frequency were not due to subjects' awareness of compositional constraints in the location of the dots.

Encoding precision. Our final analyses examined a potential limiting factor in how well subjects can estimate element-location frequency. That is, if the position codes associated with each element-level memory unit are inaccurate, the accuracy of frequency estimation will decrease. This hypothesis is based on the assumption that frequency estimates at a location depend on all the element-level memory units with position codes at the location being tested, including memory units for elements from surrounding locations whose positions were incorrectly coded. The effect, on estimation accuracy, of miscoded elements from surrounding locations will depend on the frequency with which elements occur in those locations. For example, if a location with a low frequency of occurrence is surrounded by locations with high frequencies of occurrence, imprecise position coding will result in frequency estimates for that location being too high.

To test for effects of imprecise position coding, we summed, for each of the 16 locations in the 4x4 matrix, the frequencies of occurrence for all the adjacent locations, including diagonally adjacent locations (e.g., the total adjacent-surround frequency for the location in the lower-right corner of Figure 1 is 32). We then computed what we called the effective frequency for the location. This was its actual frequency plus k (a variable weighting factor) times the total frequency of the location's "adjacent-surround." Since we had no preconceptions concerning an appropriate value for k, we computed the correlation between mean-estimated frequency and the hypothetical effective frequency for values of k ranging from zero to one. The results of these computations are presented in Figure 2. In evaluating these results, we were looking for an increase in the size of the correlation relative to that obtained when k was zero (i.e., when the adjacent-surround was not taken into account).

```
--------------------------
Insert Figure 2 about here
--------------------------
```

As indicated in Figure 2, when k = 0 the correlations based on mean-estimated frequency were 0.72 for the Vertical/Horizontal condition and 0.75 for the Diagonal condition. Both of these correlations were substantially higher than the means of the individually computed correlations. For the Vertical/Horizontal condition, the size of the correlation between mean-estimated frequency and the effective frequency (for all 16 locations) increased from 0.72, when k was zero, to a maximum of 0.90 for k = .3 (see Figure 2). A smaller increase in over-all estimation accuracy was observed for the Diagonal condition. The size of the correlation between mean-estimated frequency and the

effective frequency (for all 16 locations) increased from 0.75, when k was zero, to 0.82 for k = .2 (see Figure 2).

The above result indicated that the adjacent-surround influenced subjects' frequency estimates. While this could be attributed to the imprecision of position coding for memory units corresponding to individual elements, another possibility is that the occurrence of elements in surrounding locations influenced estimation accuracy by introducing confusions in retrieval for elements with similar position codes. This could occur if element-location was encoded with sufficient precision for the horizontal or vertical component of the memory code for one element to be the same as the horizontal or vertical component of the memory code for an element at another location. In order to test for whether estimates of element-location frequency were influenced by confusions in retrieving position codes for element-level memory units, we computed the "rectilinear-surround" for each location by summing the frequencies of element occurrence for all the locations to its left, its right, above it, and below it (e.g., the total rectilinear-surround frequency for the location in the lower-right corner of Figure 1 is 64). Our computation of effective frequency thereby included the frequencies for all locations with the same horizontal or vertical memory code as the location being estimated.

The hypothetical effective frequency for each location, its actual frequency plus k times the total frequency of the element's rectilinear-surround, was correlated with mean-estimated frequency for values of k ranging from zero to one. The results of these computations are presented in Figure 2. For the Vertical/Horizontal condition, the increase in the size of the correlation between mean-estimated frequency and the effective frequency (for all 16 locations) was smaller than that observed when the computation of effective frequency was based on the adjacent-surround. This indicated that the accuracy with which subjects estimated element-location frequency in the Vertical/Horizontal condition was limited by the imprecision of the position codes associated with traces of the individual circles rather than confusion in the retrieval of relatively precise position codes. For the Diagonal condition, the contrast between the influence of the adjacent-surround and the rectilinear-surround was smaller, but in the same direction as that observed in the Vertical/Horizontal condition.

Discussion

Prior to the frequency estimation phase of the experiment, subjects did not anticipate that they would be asked to estimate element-location frequency. When they were instructed that they would be required to estimate how often circles appeared at various locations within the frame, virtually every subject protested that they would be unable to perform the estimation task. Similar protests were received during the experiments reported later in this paper (the unexpected nature of the test was, with the exception of two subjects, always confirmed by post-experimental interviews). Subjects were assured that people always do much better than they expect at this task and were encouraged to try hard and give their best guess. Their success at estimating element-

location frequency, though modest, was obtained despite their certainty
that they lacked the knowledge to perform the estimation task.

The results of the experiment supported our hypothesis that
subjects could unintentionally encode the positions of the constituent
elements of patterns. These results were obtained despite the
likelihood that at least two performance factors limited the accuracy of
subjects' estimates. First, we suspect that many subjects were
overwhelmed by their "feeling of not knowing" and simply responded
randomly in the estimation task.[4] Second, generating a numerical
frequency estimate on the basis of element position information retrieved
from memory would itself be subject to uncertainty. Since the
correlation between actual and estimated frequency was based on only 16
data points, the results observed for individual subjects were probably
highly sensitive to performance-limiting factors like those described
above. When estimates of element-location frequency were averaged over
all the subjects participating in a condition, the influence of these
performance-limiting factors was reduced, and the correlation between
mean-estimated and actual frequency was substantially larger than the
mean of the individual correlations.

Analyses based on mean-estimated frequencies indicated that an
important source of inaccuracy in subjects' estimates of
element-location frequency was the inaccuracy with which they encoded the
position of the circles. Estimates were influenced by the frequency of
element-occurrence in locations surrounding the location being estimated.
This was indicated by the increased correlations obtained when mean-
estimated frequency was correlated with effective frequency, the latter
being based on each location's actual frequency plus k (a variable
weighting factor) times the total frequency of the location's
adjacent-surround. A similar analysis based on what we have identified
as the rectilinear-surround for each location indicated that estimation
accuracy was limited by the imprecision of position codes rather than
confusion in the retrieval of relatively precise position codes. Smaller
effects of imprecise position coding were observed for the Diagonal
compared with the Vertical/Horizontal condition.

Imprecision in position coding has previously been characterized by
probabilistic models which assume that the position code associated with
an element specifies a spatial region in which it is likely that the
element was located (Kinchla & Allan, 1969; Kinchla, 1971; Wolford,
1975). Such models could readily account for the results obtained in
this experiment, as well as the experiments that follow, if it is
assumed that subjects produce estimates for each location by retrieving
element-level memory units and incrementing their estimates of
occurrence-frequency according to the probability that each of the
retrieved elements was located at the location being tested.

The conclusion that the accuracy of subjects' estimates of
element-location frequency was limited by the imprecision of the
position codes associated with individual elements assumes that
subjects' estimates were, in fact, based on memory units corresponding
to individual elements. Our analysis of subjects' recall data indicated
that this was the case. If subjects estimated the frequency of

occurrence for each location by retrieving pattern-level memory units with elements at the location being estimated, we would not have observed a significant actual/estimate correlation after the frequency with each location appeared in subjects' correct-recall protocols was "partialled out." We are not claiming that subjects did not store pattern-level information. What we are concluding is that in this experiment, pattern-level memory units played a minimal role in the estimation of element-location frequency.

Finally, the "oblique effect" observed in so many different perceptual tasks (Appelle, 1972) was also observed in our Phase 1 detection data. Response times were significantly slower when subjects looked for diagonal alignments than when they looked for vertical or horizontal alignments. Our memory results, however, did not provide the consistent vertical/horizontal advantage that is typical of perceptual tasks. More patterns were correctly recalled in the Diagonal the Vertical/Horizontal condition, and there was no apparent advantage in frequency estimation for the Vertical/Horizontal compared with the Diagonal condition.

Experiment 2

The results of Experiment 1 indicated that subjects, without intention, can encode the frequency with which various locations within a frame are occupied by the constituent elements of a series of patterns. In this experiment we again tested whether subjects could unintentionally encode element-location frequency, but with a different Phase 1 task. Estimation performance in this task was contrasted with one in which subjects were instructed to try to remember element-location frequency. The effects of intentionality (or lack of same) constitute one of Hasher and Zacks' (1979) criteria for determining whether or not information is encoded automatically. We examined intentionality in this experiment from two points of view: 1) Does it produce differences in estimation performance? and 2) Does it produce differences in how precisely subjects encode element-location information? The first question was addressed by comparing estimation accuracy in the two conditions. The second question was addressed by comparing the effects of the adjacent-surround on estimation accuracy in the two conditions.

Method

Subjects. Sixty four undergraduate students at Florida Atlantic University participated in this experiment as one way of receiving course credit in an undergraduate psychology class.

Procedure. The experimental procedure, with the exception of the tasks introduced in Phase 1, was identical to that of Experiment 1. The patterns used in this experiment were also identical to those used in Experiment 1; the 17 patterns were each presented twice. However, the dots presented inside the circles were arranged differently in this experiment. On one occurrence of a pattern, a dot was presented in the center of two of the circles. On its other occurrence a dot was

presented in the center of the other three circles. The order of this assignment of dots (two vs. three per pattern) was randomized.

For one group of 32 subjects, the Phase 1 task for each pattern involved counting the number of circles with a dot inside it. Subjects in the Counting condition were required to press one button if there were three circles with dots and the other button if there were two circles with dots. They were instructed to respond as quickly as possible while keeping their errors to a minimum, and were given no indication that they would have to remember anything about the patterns. The 32 subjects in the Intentional condition saw the same patterns (and dots inside the circles) as in the Counting condition, but they did not have a discrimination task. They were shown an index card with a frame and grid lines defining 16 locations within the frame, and were instructed to try to remember the number of times a circle appeared in each of the locations within the frame (recall that the grid lines were never actually presented during the first two phases of the experiment).

Results

Performance in orienting tasks. Mean response times and error rates for the Counting task are presented in Table 1. Faster responses were obtained for "2" compared with "3" responses. Over-all response times and errors were similar to those obtained in the Vertical/Horizontal condition of Experiment 1.

Accuracy of frequency estimation. Correlations between the actual and estimated frequencies were again relatively small, but as indicated in Table 2, they were positive for most of the subjects in the Counting and Intentional conditions. The means of the actual/estimate correlations (see Table 2) were significantly greater than zero [$r = 0.28$, $t(31) = 5.22$, $p < .001$; $r = 0.35$, $t(31) = 8.07$, $p < .001$], for the Counting and Intentional conditions, respectively. The correlations obtained for the two conditions were not significantly different from each other, [$t(62) = 1.14$, $p > .05$].

The derivational strategy. As in Experiment 1, mean partial-correlation coefficients were somewhat reduced from the correlation coefficients observed without "partialling out" recall-frequency, but as indicated in Table 2, correlations remained positive for most of the subjects. The means of the partial correlations were significantly greater than zero [$r = 0.26$, $t(31) = 4.59$, $p < .001$; $r = 0.28$, $t(31) = 4.47$, $p < .001$], for the Counting and Intentional conditions, respectively. The partial correlations obtained for the Counting and Intentional conditions were not significantly different from each other, [$t(62) < 1.0$]. As in Experiment 1, there may have been some tendency for subjects to have derived their frequency of occurrence estimates for each location from the number of patterns they could recall with a circle at that location, but a significant actual/estimate correlation remained when the contributions of the derivational strategy were "partialled out."

Recall accuracy. As indicated in Table 3, over-all recall was more accurate in the Intentional condition (14.2% of the patterns were correctly recalled) than in the Counting condition (8.5% of the patterns

were correctly recalled). This difference was significant [t(62) = 2.00, p < .05]. The correctly recalled patterns were placed in the correct location in the 4x4 frame at a better than chance rate in the Intentional, but not the Counting condition.

Compositional constraints. The three correlations between mean-estimated and actual frequency in the Counting condition were all positive: 0.95, 0.75, and 0.31 for the corner, side, and middle locations, respectively. These correlations were also positive in the Intentional condition: 0.82, 0.73, and 0.70 for the corner, side, and middle locations, respectively. We could again conclude that the observed correlations between actual and estimated frequency were not due to subjects' awareness of compositional constraints in the location of the circles.

Insert Figure 3 about here

Encoding precision. The correlations between mean-estimated frequency and the effective frequency for each location (determined by accounting for the frequency of surrounding locations) are presented in Figure 3. As indicated in the figure, when the weighting factor, k, was equal to zero, the correlations based on mean-estimated frequency were 0.73 for the Counting condition and 0.80 for the Intentional condition. Both of these correlations were substantially higher than the means of the individually computed correlations. For the Counting condition, the correlation between mean-estimated frequency and effective frequency (based on the adjacent-surround) increased from 0.73, when k was zero, to 0.96 for k = .6. A smaller increase was obtained when effective frequency was based on the rectilinear-surround: the correlation between mean-estimated frequency and effective frequency increased from 0.73, when k was zero, to 0.81 for k = .2. Estimation accuracy appeared to be influenced by imprecise position coding in the Counting condition.

In the Intentional condition, basing our correlations on effective rather than actual correlations resulted in a relatively small improvement in estimation accuracy when effective frequency was based on the adjacent-surround (from r = 0.80 when k = 0, to r = 0.85 when k = .2), and a sharp decrease in estimation accuracy when effective frequency was based on the rectilinear-surround. Imprecise position coding had a smaller effect on estimation accuracy in the Intentional condition compared with the unintentional, Counting condition.

Discussion

The results of this experiment replicated all the observations made in Experiment 1. Subjects were able to estimate the frequency with which circles appeared at various locations in the frame, they did not appear to derive their estimates from pattern-level memory units or their knowledge of compositional constraints regarding the likely location of circles in the frame, and uncertainty in position coding was the result of imprecise position coding rather than confusions in retrieval. Although the mean of the individual actual/estimate

correlations and the mean-estimate/actual correlation were slightly better in the Intentional than the Counting condition, the differences were not significant. These results did not allow us to draw any conclusions regarding Hasher and Zacks' (1979) criterion for automatic encoding. However, further analyses indicated that the frequency with which circles appeared in the adjacent-surround of each location had a greater effect on frequency estimation in the Counting compared with the Intentional condition. Thus, element-position appeared to be coded more precisely in the Intentional compared with the Counting condition.

Experiment 3

In the previous two experiments we obtained unanticipated differences in the accuracy of pattern recall. In Experiment 1, recall was more accurate in the Diagonal condition than in the Vertical/Horizontal condition. In Experiment 2, recall was more accurate in the Intentional than in the Counting condition. What made these differences potentially important was that they appeared to be related to how precisely subjects encoded the position of individual circles (as indicated by effects of the adjacent-surround on estimation accuracy). That is, relatively high levels of pattern recall appeared to be associated with relatively precise encoding of position, whereas relatively low levels of pattern recall appeared to be associated with relatively imprecise encoding of position. Although this suggested that pattern-level processing facilitates the encoding of element-level position information, better recall and more precise location coding did not result in more accurate frequency estimation.

The purpose of this experiment was to clarify the relationship between pattern-level processing and the encoding of element position information. The Phase 1 tasks used in this experiment were selected with the anticipation that they would be maximally contrastive in their effects on recall accuracy. In one condition, element-level processing was emphasized by the Phase 1 task; subjects were required to detect the present of a target (a small dot) inside one of the five circles comprising each pattern. In the second condition, pattern-level processing was emphasized by the Phase 1 task; subjects were told to try to remember each of the patterns. As in the previous experiments, we compared the groups assigned to the two conditions with regard to their estimation accuracy and the precision with which they coded element position. The comparison allowed us to determine whether the formation of pattern-level memory units is compatible with, or competes with the formation of element-level memory units.

Method

Subjects. Sixty four undergraduate students at Florida Atlantic University participated in this experiment as one way of receiving course credit in an undergraduate psychology class.

Procedure. The stimuli and experimental procedure, with the exception of the tasks introduced in Phase 1, were identical to that of Experiments 1 and 2. However, the dots presented inside the circles were arranged in accordance with the Phase 1 tasks used in this

experiment. On one occurrence of a pattern, a dot was presented in the center of one of the circles, on the other occurrence of the pattern, a dot was not presented. The order of this assignment of dots (zero vs. one per pattern) was randomized.

For one group of 32 subjects, the Phase 1 task for each pattern involved detecting a circle with a dot inside it. Subjects in the Present/Absent condition were required to press one button if there was a dot present and the other button in the absence of a dot. They were instructed to respond as quickly as possible while keeping their errors to a minimum, and were given no indication that they would have to remember anything about the patterns. The 32 subjects in the Pattern Memory condition saw the same patterns (and dots inside the circles) as in the Present/Absent condition, but they did not have a discrimination task; they were instructed to try to remember each pattern. However, they were not told that they would subsequently be asked to estimate the number of times a circle appeared in each of the locations within the frame.

Results

Performance in orienting tasks. Mean response times and error rates are presented in Table 1. There was little difference in performance for Present and Absent responses.

Accuracy of frequency estimation. As in the previous two experiments, correlations between the actual and estimated frequencies for each location in the 4x4 matrix were relatively small, but as indicated in Table 2, they were positive for most of the subjects in the Present/Absent and Pattern Memory conditions. The means of the actual/estimate correlations (see Table 2) were significantly greater than zero [$r = 0.35$, $t(31) = 7.64$, $p < .001$; $r = 0.17$, $t(31) = 3.24$, $p < .02$], for the Present/Absent and Pattern Memory conditions, respectively. The correlations obtained for the two conditions were significantly different from each other, [$t(62) = 2.55$, $p < .05$].

The derivational strategy. In the Present/Absent condition, partial-correlations remained positive for most of the subjects and the mean partial-correlation coefficient was only slightly reduced from the mean obtained without partialling out recall-frequency (see Table 2). The mean of the partial correlations was significantly greater than zero [$r = 0.33$, $t(31) = 7.24$, $p < .001$]. In the Pattern Memory condition, however, partial correlations obtained for individual subjects were evenly divided between positive and negative values, and the mean partial-correlation coefficient was reduced to the point where it was no longer significantly different from zero [$r = 0.03$, $t(31) < 1.0$]. The partial correlations obtained for the Present/Absent and Pattern Memory conditions were significantly different from each other, [$t(62) = 4.06$, $p < .001$]. In contrast with the Present/Absent condition, it was likely that subjects in the Pattern Memory condition derived their frequency of occurrence estimates for each location from the number of patterns they could recall with a circle at that location.

Recall accuracy. As indicated in Table 3, over-all recall was more accurate in the Pattern Memory condition (17.8% of the patterns were correctly recalled) than in the Present/Absent condition (7.4% of the patterns were correctly recalled). This difference was significant [$t(62)$ = 5.17, $p < .001$]. In both conditions, the correctly recalled patterns were placed in the correct location in the 4x4 frame at a better than chance rate.

Compositional constraints. Large correlations between mean-estimated and actual frequency in the Present/Absent condition were obtained for the corner (r = 0.97) and side locations (r = 0.87), but not the middle locations (r = 0.10). All three correlations were strongly positive in the Pattern Memory condition: 0.99, 0.66, and 0.55 for the corner, side, and middle locations, respectively. We could again conclude that the observed correlations between actual and estimated frequency were not due to subjects' awareness of compositional constraints in the location of the dots.

Insert Figure 4 about here

Encoding precision. For the Present/Absent condition (see Figure 4), the correlation between mean-estimated frequency and effective frequency (based on the adjacent-surround) increased from 0.76, when k was zero, to 0.93 for k = .3. A smaller increase was obtained when effective frequency was based on the rectilinear-surround: the correlation between mean-estimated frequency and effective frequency increased from 0.76, when k was zero, to 0.89 for k = .3. For the Pattern Memory condition, computing the effective frequency of each location did not increase estimation accuracy when effective frequency was based on the adjacent-surround, and only slightly increased estimation accuracy when effective frequency was based on the rectilinear-surround. These results indicated, therefore, that element position was coded more precisely in the Pattern Memory than in the Present/Absent condition.

Discussion

The results were consistent with those observed in the previous two experiments. That is, relatively high levels of pattern recall appeared to be associated with relatively precise encoding of position (the Pattern Memory condition), whereas relatively low levels of pattern recall appeared to be associated with relatively imprecise encoding of position (the Present/Absent condition). However, high levels of recall and relatively precise position coding were accompanied by low levels of estimation accuracy. That is, subjects in the Pattern Memory condition estimated element-location frequency with significantly less accuracy than subjects in the Present/Absent condition. Furthermore, the limited ability of subjects in the Pattern Memory condition to estimate element-location frequency could be accounted for, almost entirely, by the retrieval of pattern-level memory units. When we counted the frequency with which each location was occupied by a circle in each subject's correct-recall protocols, and computed a partial-correlation

coefficient for each subject in which the recall-frequency for each location was "partialled out," the correlation between the actual and estimated frequency was reduced to nonsignficance for the subjects in the Pattern Memory condition. Since this was not the case in the Present/Absent condition, we could conclude that subjects' estimates in the Present/Absent condition were based on element-level memory units.

Experiment 3 therefore indicated that: 1) the derivation of frequency estimates from the retrieval of pattern-level memory units resulted in poor estimation of element-location frequency, and 2) emphasis on the formation of pattern-level memory units suppressed the formation of element-level memory units that are the basis for the relatively accurate frequency estimation. This was the case even though element position was encoded more precisely for pattern-level memory units than for element-level memory units. Given their importance, the purpose of Experiment 4 was to replicate the results obtained in Experiment 3.

Experiment 4

Method

Subjects. Sixty four undergraduate students at Florida Atlantic University participated in this experiment as one way of receiving course credit in an undergraduate psychology class.

Procedure. The only difference in procedure compared with Experiment 3 concerned the nature of the Phase 1 detection task. In this experiment, a dot appeared in one circle for every pattern presented during Phase 1. The dot was displaced to either the right or left of center. It was presented in different circles on the two presentations of each pattern, once displaced to the right and once displaced to the left.[5] The order of this assignment of dot-displacements (right vs. left) was randomized.

The group of 32 subjects in the Right/Left condition was required to press one button if the dot present inside one circle was displaced to the right and the other button if it was displaced to the left. They were instructed to respond as quickly as possible while keeping their errors to a minimum, and were given no indication that they would have to remember anything about the patterns. The 32 subjects in the Pattern Memory condition saw the same patterns (and dots inside the circles) as in the Right/Left condition, but they did not have a discrimination task. They were instructed to try to remember each pattern and were not told that they would subsequently be asked to estimate the number of times a circle appeared in each of the locations within the frame.

Results

Performance in orienting task. Mean response times and error rates for the Right/Left judgment task are presented in Table 1. "Right" responses were somewhat faster than "Left" responses; error rates were slightly lower for "Left" responses.

Accuracy of frequency estimation. Correlations between the actual
and estimated frequencies for each location in the 4x4 matrix were again
relatively small, but as indicated in Table 2, they were positive for
most of the subjects in the Right/Left and Pattern Memory conditions.
The mean of the actual/estimate correlations (see Table 2) was
significantly greater than zero [r = 0.31, t(31) = 5.32, p < .001; r =
0.19, t(31) = 3.20, p < .02], for the Right/Left and Pattern Memory
conditions, respectively. The correlations obtained for the two
conditions were not significantly different from each other, [t(62) =
1.47, p > .05]. The latter result did not replicate the comparable
difference observed in Experiment 3.

The derivational strategy. In the Right/Left condition, partial
correlations remained positive for most of the subjects and the mean
partial-correlation coefficient was only slightly reduced from the mean
obtained without "partialling out" recall-frequency (see Table 2). The
mean partial correlation was significantly greater than zero [r = 0.28,
t(31) = 4.63, p < .001]. However, in the Pattern Memory condition, the
partial correlations obtained for individual subjects were again evenly
divided between positive and negative values, and the mean partial-
correlation coefficient was reduced to the point where it was no longer
significantly different from zero [r = 0.10, t(31) = 1.68, p > .05]. The
partial correlations obtained for the Right/Left and Pattern Memory
conditions were significantly different from each other, [t(62) = 2.30, p
< .05]. As in Experiment 3, it was likely that subjects in the Pattern
Memory condition derived their frequency of occurrence estimates for each
location from the number of patterns they could recall with a circle at
that location.

Recall accuracy. Consistent with the results of Experiment 3 (see
Table 3), over-all recall was more accurate in the Pattern Memory
condition (19.7% of the patterns were correctly recalled) than in the
Right/Left condition (6.4% of the patterns were correctly recalled).
This difference was significant [t(62) = 5.83, p < .001]. In both
conditions, the correctly recalled patterns were placed in the correct
location within the 4x4 frame at a better than chance rate.

Compositional constraints. In the Right/Left condition, positive
correlations between mean-estimated and actual frequency were obtained
for the corner (r = 0.73) and side locations (r = 0.47), but the
correlation was negative for the middle locations (r = -0.57). However,
all three correlations were positive in the Pattern Memory condition:
0.75, 0.47, and 0.73 for the corner, side, and middle locations,
respectively. Although the evidence was not quite as strong as in the
previous three experiments, we could again conclude that subjects'
frequency estimates were not based on their awareness of compositional
constraints regarding where circles were most likely to be located.

Insert Figure 5 about here

Encoding precision. The results of analyzing the precision with
which subjects encoded element position were quite similar to those

obtained in Experiment 3. For the Right/Left condition (see Figure 5), the correlation between mean-estimated frequency and effective frequency (based on the adjacent-surround) increased from 0.61, when k was zero, to 0.90 for k = .7. A much smaller increase was obtained when effective frequency was based on the rectilinear-surround: the correlation between mean-estimated frequency and effective frequency increased from 0.61, when k was zero, to 0.76 for k = .4. For the Pattern Memory condition, computing the effective frequency of each location did not increase estimation accuracy when effective frequency was based on either the adjacent- or rectilinear-surround. The results indicated, once again, that element position was coded more precisely in the Pattern Memory than in the Right/Left condition.

Discussion

The results of this experiment replicated those of Experiment 3 in most respects. As in Experiment 3, the Phase 1 task that emphasized pattern-level processing (the Pattern Memory condition) resulted in better recall and more precise element-position coding than the Phase 1 task that emphasized element-level processing (the Right/Left condition). Also as in Experiment 3, subjects in the Pattern Memory condition appeared to base their estimates of element-location frequency on the retrieval of pattern-level memory units. In contrast, subjects in the Right/Left condition based their frequency estimates on the retrieval of element-level memory units. Subjects in the Pattern Memory condition were again less accurate in estimating frequency than subjects receiving an element-level task during Phase 1 (the Right/Left condition), but the difference was not statistically significant in this experiment.

General Discussion

Despite their certainty that they lacked the knowledge to perform the estimation task, when coaxed to do so, subjects were able to estimate the frequency with which the constituent elements of a series of patterns occurred in various locations within a frame. This result was consistent with other evidence indicative of a dissociation between what subjects know and what they think they know. For example, Graf, Mandler, and Haden (1982) had their subjects look for vowel repetitions in successively presented words. Although they were unable to recall these words, their completion of word stems was biased by the previously seen words that they could not recall. Evidence of a dissociation between what subjects know and what they think they know is most vividly observed with various forms of brain damage. Amnesics show evidence of retention for material they do not recall learning (see Parkin [1982] for a review) and people who report that they have no visual experience in one visual field, when coaxed, can successfully point to targets in their "blind" field (Weiskrantz, Warrington, Sanders, & Marshall, 1974).

As indicated earlier, it was possible for our subjects to have produced reasonably accurate estimates of element-location frequency without remembering anything about the patterns. They could have done so if they realized that there were compositional constraints which influenced where the circles composing our patterns were most likely to

be located; the circles were constrained to occur in the middle locations
most often and the corner locations least often. However, awareness of
these constraints was not sufficient for subjects to have differentiated
among frequencies of element occurrence at each of the four corner
locations, at each of eight side locations, or at each of the four
middle locations. In all four experiments, we obtained evidence from
subjects' mean-estimates indicating that they could independently
estimate frequency at the corner, side, and middle locations. We
concluded, therefore, that subjects' estimates were based, not on their
awareness of compositional constraints in the location of the circles,
but on what they remembered about the stimuli presented during Phase 1.

The experimental results indicated that accurate estimation of
element occurrence frequency depended on the formation of element-level
memory units. When the Phase 1 task emphasized element-level processing
(the Present/Absent and Right/Left conditions of Experiments 3 and 4),
pattern-level memory units had little influence on subjects' estimates,
which were relatively accurate. However, estimation accuracy decreased
sharply when the Phase 1 task stressed the formation of pattern-level
rather than element-level memory units (the Pattern Memory conditions of
Experiments 3 and 4). Thus, the formation of pattern-level memory units
did not support accurate frequency estimation, and the emphasis on
pattern-level processing suppressed the formation of the element-level
memory units that would have supported accurate frequency estimation.

Evidence of reduced estimation accuracy as a result of pattern-
level processing was surprising for two reasons: 1) pattern-level
processing requires attending to all the elements of each pattern; the
element-level tasks could be accomplished without attending to all the
elements, and 2) evidence accumulated over all four experiments indicated
that the formation of pattern-level memory units was associated with the
relatively precise encoding of element position. This association is
indicated by the scatterplot presented in Figure 6. The eight points on
the scatterplot represent the two conditions in each of the four
experiments. The location of each point is determined jointly by mean
recall accuracy (which reflects the formation of pattern-level memory
units) and the proportion of the variance in subjects' mean-estimates
that could be accounted for by the imprecise encoding of element
position. The determination of the latter was based on our analysis of
the effective frequency of element occurrence for each of the 16
locations within the frame. The effective frequency computed for each
location was the actual frequency for the location plus a variable
weighting factor (k) times the sum of the frequencies in the locations
surrounding the location in question. What we observed, in varying
degrees, was an increase in the accuracy of subjects' mean-estimates when
the computation of effective frequency included the adjacent-surround.
Since comparable increases were not obtained when effective frequency was
based on the rectilinear-surround rather than the adjacent-surround, we
concluded that the estimation data were influenced by imprecision in
position coding rather than confusions in the retrieval of relatively
precise position codes.

Insert Figure 6 about here

For all the experimental conditions showing increased correlations when effective frequency was computed on the basis of the adjacent-surround, correlations with mean-estimated frequency were at or close to their maximum when k = .3. The data in Figure 6 are therefore based on this value of k. The proportion of the variance in subjects' mean-estimates that could be accounted for by the imprecise encoding of element position was the difference in the squared correlations obtained for k = .3 (when the effect of the adjacent-surround was maximized) and k = 0 (when the effect of the adjacent-surround was not taken into account). The scatterplot in Figure 6 indicates that there was a strong negative correlation (r = -0.96) between pattern recall accuracy and the amount of variance attributable to imprecise element position coding. That is, the formation of pattern-level memory units was associated with relatively precise coding of element position.

The association between accurate pattern-recall and precise position coding suggests an important role for the perception of spatial relations among the elements of each pattern. The formation of pattern-level memory units requires encoding spatial relations among each pattern's constituent elements, and the precise determination of each element's position could result from the constraints imposed by the encoding of multiple spatial relations involving the element (e.g., an element might be alongside a nearby element, diagonally below a second element, and relatively far from a third element directly above it). Despite their precision, the element-relative position codes for pattern-level memory units will be useless for estimating element-location frequency if the global location of the patterns, relative to the frame, has not also been encoded. The results summarized in Table 3 indicate that correctly-recalled patterns were usually recalled in the correct quadrant of the frame at a greater than chance rate, but subjects were not highly accurate in this regard; the maximum proportion of correctly recalled patterns that were recalled in the correct quadrant was .68 in the Pattern Memory condition of Experiment 3. Hence, the poorest frequency estimation was obtained for the experimental conditions in which position coding was most precise (the Pattern Memory conditions) because the usefulness of the element-relative position codes associated with pattern-level memory units was limited by the relatively inaccurate encoding of each pattern's global location within the frame.

When subjects in the Intentional condition of Experiment 2 tried to remember the frequency with which circles occupied the various locations within the frame, their estimation accuracy was only marginally better than that obtained under unintentional conditions. More interesting was the indication that estimation performance in the Intentional condition was based primarily on element-level memory units (correlations between actual and estimated frequency remained significant even after the recall-frequency for each location was "partialled out"), and the position codes for the element-level memory units were more precise than in the unintentional, Counting condition of Experiment 2 (as indicated by the effects of the adjacent-surround on estimation performance). These results suggest that subjects trying to remember element-location frequency directed at least some effort to processing spatial relations

among the elements within each pattern. This increased the precision of
frame-relative position coding for the element-level memory units,[6] and
as indicated by the recall data, enhanced the formation of pattern-level
memory units (which requires encoding relations among the elements).

The results obtained in the Vertical/Horizontal condition of
Experiment 1 were consistent with the conclusion that estimates were
based on element-level memory units with frame-relative position codes;
frequency estimates remained significant even after the recall-frequency
for each location was "partialled out" of correlations between estimated
and actual frequency, and a relatively large proportion of the variance
in subjects' mean-estimates was attributable to imprecision in position
coding. In retrospect, it appears that there was little processing of
spatial relations among the elements of each pattern, perhaps because
judgments of vertical/horizontal alignment were based on whether there
were three circles parallel to the vertical or horizontal sides of the
frame. Since the opportunity to use the sides of the frame as the basis
of alignment judgment was not available in the Diagonal condition,
subjects in that condition had to base their alignment judgments on the
relative location of the elements. The need to process spatial
relations among the elements would account for the higher level of
pattern-recall accuracy (the formation of pattern-level memory units
requires the coding of element-relative position), and the more precise
position coding in the Diagonal compared with the Vertical/Horizontal
condition.

To summarize, the processing of spatial relations among the elements
of each pattern enhanced the formation of pattern-level memory units and
increased the accuracy of element position coding for both element-level
and pattern-level memory units. However, estimation accuracy suffered
for Phase 1 orienting tasks that stressed the formation of pattern-level
memory units because: 1) emphasis on the formation of pattern-level
memory units suppressed the formation of element-level memory units, and
2) despite the precision of their element position codes, imprecise
coding of global location limited the usefulness of pattern-level memory
units for the estimation of element-location frequency. Frequency
estimation in this study was best supported by element-level memory units
with frame-relative position codes, even though the orienting tasks which
emphasized element-level processing could be performed without attending
to all the elements in each pattern.

The frame-relative position of the constituent elements of a pattern
would, by most standards, be considered highly superficial information
for subjects to be retaining. Our results therefore join evidence based
on exact repetition effects (e.g., Jacoby & Brooks, 1984; Pollatsek, et
al., 1984; Hock, et al., 1981; Kolers, 1976) in providing evidence for
the retention of meaningless, superficial information. One way of
explaining the retention of superficial details for patterns is to
specify that patterns are stored and retrieved as literal, template-like
pictorial copies. However, it has been argued (Neisser, 1967; Dodwell,
1970) that template models lack the flexibility required for stimulus
generalization: the ability to recognize altered versions of previously
seen patterns (e.g., Attneave, 1957). The results reported in this paper
indicate that superficial element-position information does not have to

be retained at the pattern-level in the form of template-like pictorial copies. Superficial details involving element position were abstracted from a series of patterns and stored in the form of element-level memory units in the same way that semantically important, shared attributes might be abstracted from a series of patterns belonging to the same category (e.g., Posner & Keele, 1968). There is no reason, therefore, to assume that superficial details are represented any differently than the semantically important attributes that are the basis for stimulus generalization.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique" effect in man and animals. Psychological Bulletin, 78, 266-278.

Attneave, F. (1957). Transfer of experience with a class-schema to identification-learning of patterns and shapes. Journal of Experimental Psychology, 54, 81-88.

Bell, H. H., & Handel, S. (1976). The role of pattern goodness in the reproduction of backward masked patterns. Journal of Experimental Psychology: Human Perception and Performance, 2, 139-150.

Clement, D. E., & Weiman, C. F. R. (1970). Instructions, strategies and pattern uncertainty in a visual discrimination task. Perception & Psychophysics, 7, 333-336.

Dodwell, P. C. (1970). Visual pattern recognition. New York: Holt, Rinehart and Winston.

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. Journal of Verbal Learning and Verbal Behavior, 2, 446-452.

Graf, P., Mandler, G., & Haden, P. E. (1982). Simulating amnesic symptoms in normal subjects. Science, 218, 1243-1244.

Hasher, L. & Zacks, R. T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 108, 356-388.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. Journal of Experimental Psychology, 88, 297-306.

Hock, H. S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. Journal of Experimental Psychology: Learning, Memory & Cognition, 12, 232-240.

Hock, H. S., Throckmorton, B., Webb, E., & Rosenthal, A. (1981). The effects of graphemic and phonemic processing on the retention of words and nonwords. Memory & Cognition, 9, 461-471.

Jacoby, L. L., & Brooks, L. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.) The psychology of learning and motivation: Advances in research and theory, (Vol. 18). New York: Academic Press.

Kinchla, R. A. (1971). Visual movement discrimination: A comparison of absolute and relative movement discrimination. Perception & Psychophysics, 9, 165-171.

Kinchla, R. A., & Allan, L. G. (1969). A theory of visual movement perception. Psychological Review, 76, 537-558.

Kolers, P. A. (1976). Pattern-analyzing memory. Science, 191, 1280-1281.

Miller, J. (1981). Global precedence in attention and decision. Journal of Experimental Psychology: Human Perception and Performance, 7, 1161-1174.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 9, 353-383.

Neisser, U. (1967). Cognitive psychology. Englewood Cliffs, New Jersey: Prentice-Hall.

Parkin, A. J. (1982). Residual learning capability in organic amnesia. Cortex, 28, 417-440.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. Journal of Experimental Psychology: General, 113, 426-442.

Pomerantz, J. R., & Sager, L. C. (1975). Asymmetric integrality with dimensions of visual pattern. Perception & Psychophysics, 18, 460-466.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8, 382-439.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. Brain, 97, 709-728.

Wolford, G. (1975). Perturbation model for letter detection. Psychological Review, 82, 184-199.

Author Notes

Footnotes

[1]Because continuous contours could not be displayed on the monitor, the elements constituting each pattern were approximations of circles.

[2]Individually computed correlation coefficients served as descriptive statistics in this study. The computation of mean correlation coefficients and t-tests were based on Fischer's r to z transformation.

[3]The 17 patterns used in this study were classifiable according to the size of their equivalence set, which is determined by the number of patterns that can be generated by rotations and reflections of the pattern with respect to its horizontal and vertical axes (Garner & Clement, 1963). There were two patterns with equivalence sets of size 1, eight with equivalence sets of size 4, and seven with equivalence sets of size 8. All four experiments reported in this study replicated Bell and Handel's (1976) evidence that the recall accuracy for a pattern was inversely related to the size of its equivalence set. The results were consistent with the interpretation of equivalence set size as being related to the configural complexity or "goodness of form" of a pattern (Garner & Clement, 1963).

[4]Over all four experiments, four of the subjects were replaced because they repeated the same estimate for every location tested.

[5]In order to be certain that subjects' estimates of element-location frequency depended on the frequency of occurrence of circles rather than the frequency of occurrence of the dots placed inside some of the circles, we selected dot locations in this experiment so that there was no relationship between circle and dot frequency (the correlation between them was -0.09). The correlations between circle and dot frequency in some of the previous experiments turned out, by chance, to be relatively high. In order to determine whether dot frequency influenced the results obtained in these experiments, we computed partial-correlation coefficients with the effects of dot frequency "partialled out." Although there were some fluctuations in individual data, the over-all pattern of data was unchanged. This indicated that subjects' estimates of how often circles occurred at locations within the frame were indeed based on circle frequency rather than dot frequency.

[6]How the processing of spatial relations among the elements of a pattern might increase the accuracy of frame-relative position coding for individual elements is a topic for further investigation. One possibility is that codes based on the perceived distance between adjacent elements in the pattern serve as "mental units of measurement" which enhance the precision with which subjects encode the distance of each element from the sides of the frame. Another possibility is that the perception of angular relations between pairs of elements in the pattern requires encoding the frame-relative position of the elements to a greater level of precision than would otherwise be the case.

Table 1

Mean Reaction Times (in msec) and Percent Errors for the Phase 1 Orienting Tasks.

Experiment	Phase 1 Task	Response Type	Mean Reaction Time	Percent Errors
1	Vertical/Horizontal Alignment	Yes	945	6.5
		No	1056	3.5
	Diagonal Alignment	Yes	1022	7.5
		No	1070	2.3
2	Counting	"2"	932	2.0
		"3"	996	4.8
3	Present/Absent	Present	845	3.5
		Absent	865	2.2
4	Right/Left	Right	1041	6.3
		Left	1097	5.4

Table 2

Summary of the Correlation Coefficients Between Estimated Frequency (Obtained During Phase 2) and Actual Frequency for the 16 Locations Within the Frame. The Partial Correlations Were Computed by "Partialling Out" the Effect of Recall-Frequency (Determined from the Phase 3 Recall Data) from the Estimated/Actual Correlation.

Experiment	Phase 1 Task	Correlation Actual with Estimate		Partial Correlation: Actual with Estimate	
		Proportion of Subjects with Positive Correlation	Mean Correlation	Proportion of Subjects with Positive Correlation	Mean Correlation
1	Vertical/Horizontal Alignment	.78	0.34	.81	0.26
	Diagonal Alignment	.84	0.33	.88	0.28
2	Counting	.81	0.28	.78	0.24
	Intentional	.94	0.35	.75	0.30
3	Present/Absent	.92	0.35	.88	0.33
	Pattern Memory	.78	0.17	.53	0.03
4	Right/Left	.81	0.31	.81	0.28
	Pattern Memory	.63	0.19	.50	0.10

Table 3

Percentage of Patterns Correctly Recalled, Regardless of Whether They were Recalled in the Correct Location Within the Frame, and the Proportion of These Correctly Recalled Patterns That Were Recalled in the Correct Location Within the Frame.

Experiment	Phase 1 Task	Percent Correct Recall	Proportion of Correctly Recalled Patterns in Correct Quadrant
1	Vertical/Horizontal Alignment	8.1	.50
	Diagonal Alignment	12.3	.54
2	Counting	8.5	.30
	Intentional	14.2	.58
3	Present/Absent	7.4	.50
	Pattern Memory	17.8	.68
4	Right/Left	6.4	.46
	Pattern Memory	19.7	.67

Figure Captions

Figure 1. The frequency with which circles occupied each of the 16 possible locations within the frame (for the 0 deg orientation of the patterns).

Figure 2. Experiment 1: For the Vertical/Horizontal and Diagonal conditions, correlations between mean-estimated and effective frequency for the 16 locations within the frame. Each correlation was computed for a different value of k, which weights the contribution of occurrence frequencies at surrounding locations in relation to a location's actual occurrence frequency in determining the effective frequency of occurrence for the location. The solid lines are for correlations based on the adjacent-surround, the broken lines for correlations based on the rectilinear-surrround.

Figure 3. Experiment 2: For the Counting and Intentional conditions, correlations between mean-estimated and effective frequency for the 16 locations within the frame. Each correlation was computed for a different value of k, which weights the contribution of occurrence frequencies at surrounding locations in relation to the location's actual occurrence frequency in determining the effective frequency of occurrence for the location. The solid lines are for correlations based on the adjacent-surround, the broken lines for correlations based on the rectilinear-surrround.

Figure 4. Experiment 3: For the Present/Absent and Pattern Memory conditions, correlations between mean-estimated and effective frequency for the 16 locations within the frame. Each correlation was computed for a different value of k, which weights the contribution of occurrence frequencies at surrounding locations in relation to the location's actual occurrence frequency in determining the effective frequency of occurrence for the location. The solid lines are for correlations based on the adjacent-surround, the broken lines for correlations based on the rectilinear-surrround.

Figure 5. Experiment 4: For the Right/Left and Pattern Memory conditions, correlations between mean-estimated and effective frequency for the 16 locations within the frame. Each correlation was computed for a different value of k, which weights the contribution of occurrence frequencies at surrounding locations in relation to the location's actual occurrence frequency in determining the effective frequency of occurrence for the location. The solid lines are for correlations based on the adjacent-surround, the broken lines for correlations based on the rectilinear-surrround.
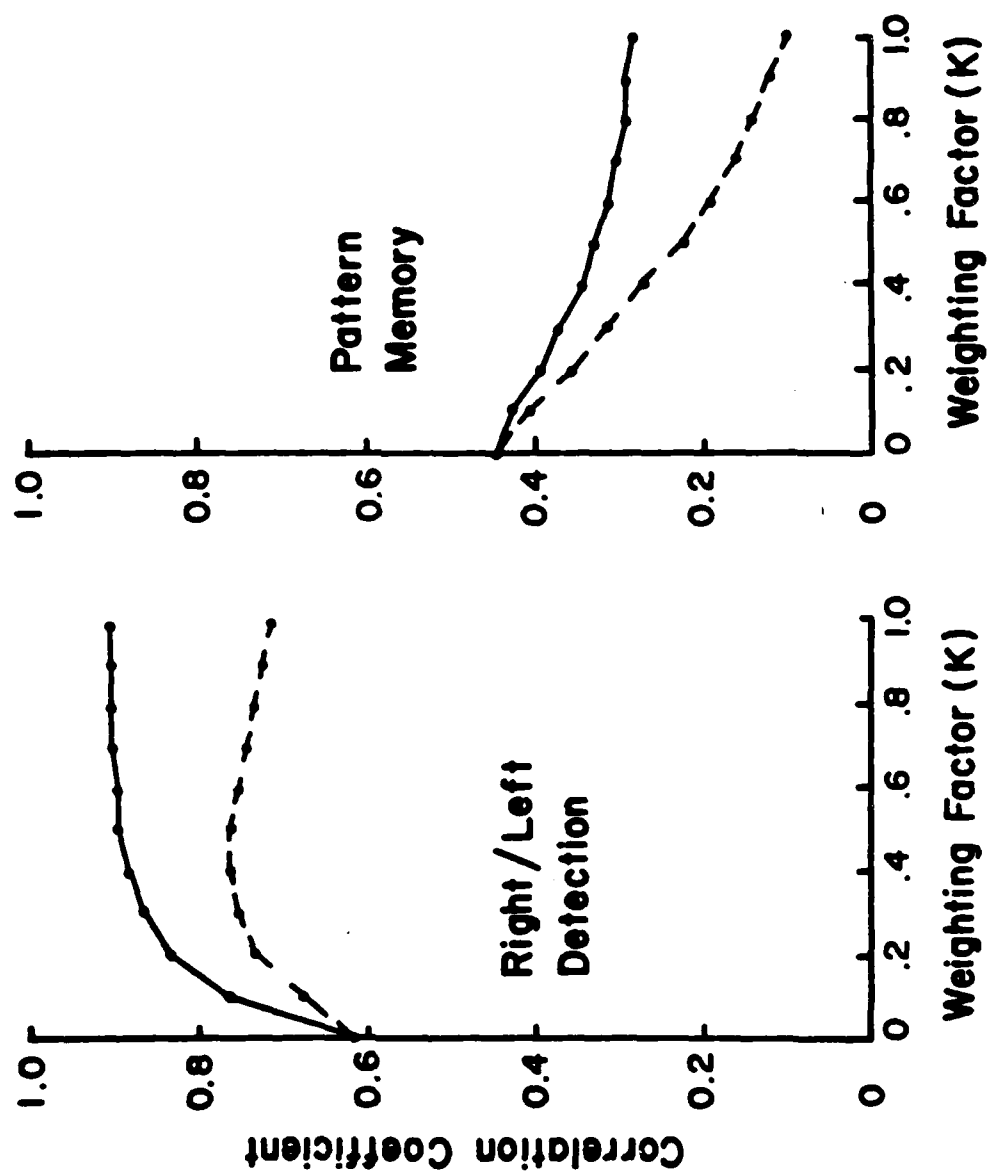
Figure 6. Scatterplot with eight points, each point representing the recall accuracy and the precision of position coding for one of the two conditions in the four experiments reported in this study.

4	2	4	4
8	22	16	10
6	6	28	8
6	20	14	12
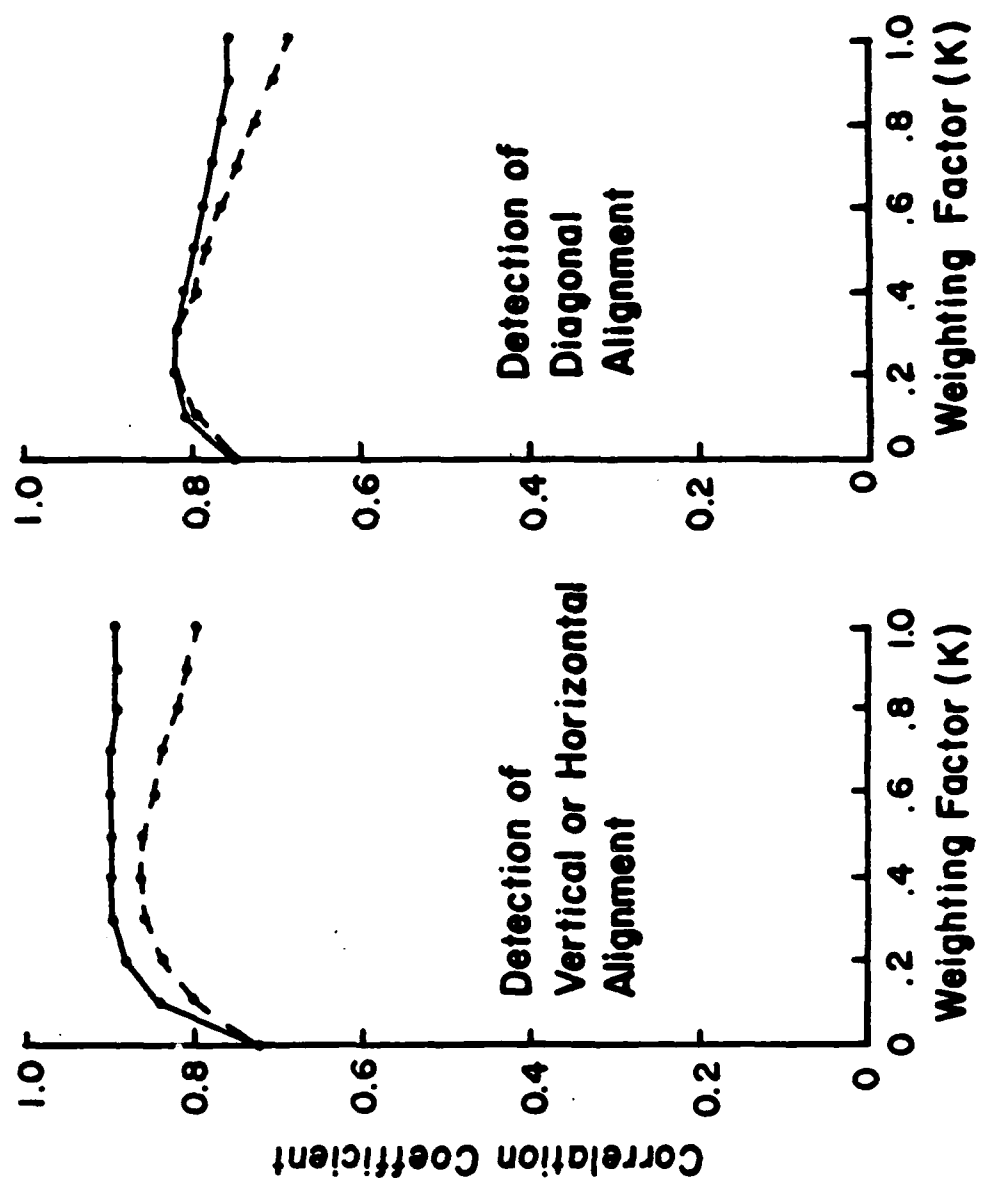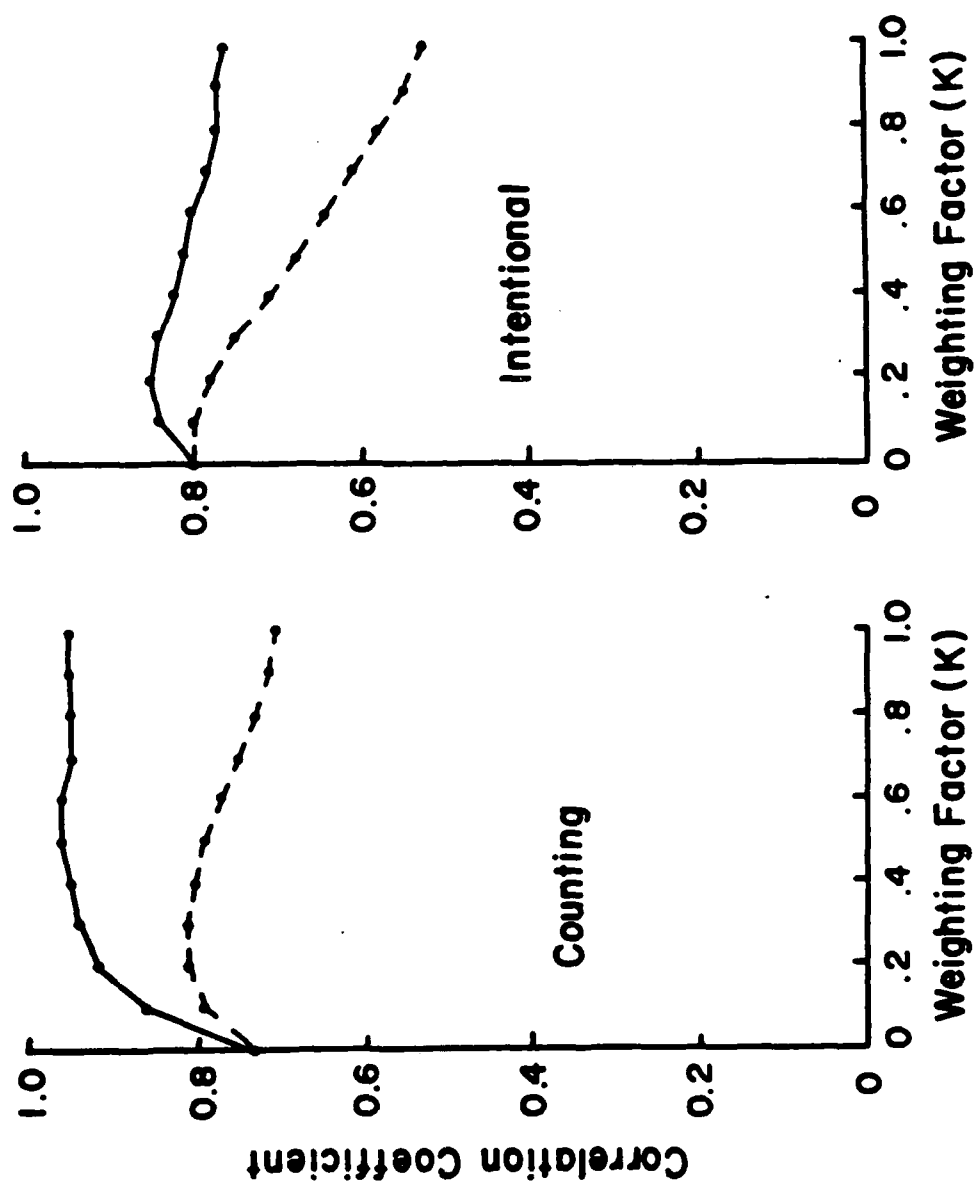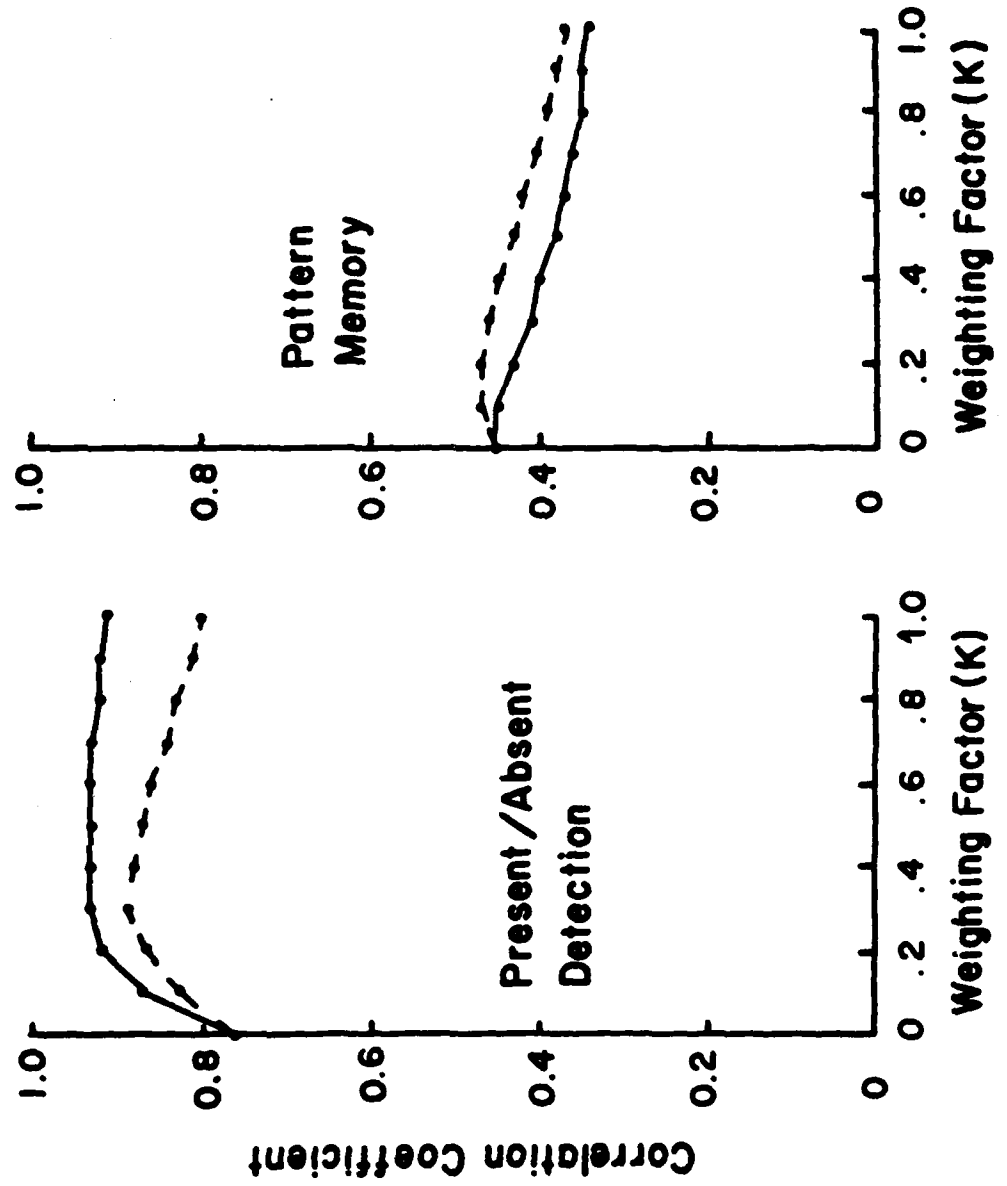
SECTION 2D

The Effect of Nonspecific Memory Instructions on the Encoding of Frequency of Occurrence Information

Howard S. Hock and L. Clayton Cavedo

Greene (1984) has recently argued that the absence of an intentional/incidental in frequency coding in the experiments reported by Hasher and Zacks (1979) may have been due to their use of nonspecific memory instructions ("try to remember the items") in the incidental condition. Such instructions, Greene argued, could result in similar frequency judgments in the intentional and incidental conditions because of the use of similar encoding strategies (e.g., covert rehearsal) in both conditions. If Greene is correct, an intentional/incidental difference in frequency judgment should emerge when the incidental "orienting task" is completely incidental.

The stimuli used to test this hypothesis were strings composed of four letters; half were common English words, half were orthographically regular nonwords. One group of subjects participated in a Lexical Decision task, which required that they discriminate between the words and nonwords. Frequency learning was completely incidental for these subjects. The second group of subjects was instructed to try and remember the information presented. Successful frequency discrimination for the Lexical group would replicate Hock, Malcus, and Hasher's (1986) evidence that the encoding of frequency information does not require intentional effort. By comparing performance for the Lexical and Memory groups, we assessed whether any advantage in the accuracy of frequency judgments could result from processing strategies elicited by nonspecific memory instructions.

Method

Design and subjects

This experiment involved a 2 (lexical vs. memory instructions) x 2 (words vs. nonwords) x 6 (frequency levels 2, 4, 6, 8, 10, 12) design. The only between group variable was "instructions." Sixteen undergraduate students at Florida Atlantic University voluntarily participated in the experiment, for which they were paid $2.00. Half the subjects were assigned to the Lexical group and half were assigned to the Memory group.

Materials

As indicated above, there were six frequency levels in this experiment. Three common words and three orthographically regular nonwords were assigned to each of the frequency levels, resulting in a total of 18 different words and 18 different nonwords. With strings repeated according to their assigned frequency level, there was a total of 252 strings in the stimulus list. The frequency level of the strings, and whether they were words or nonwords, was varied randomly in the stimulus sequence. The vast majority of repetitions were separated in the stimulus sequence by a minimum of nine strings. Only five repetitions had fewer than nine intervening strings. Half the subjects in each instructional condition were presented with one

stimulus sequence, the other half with the same sequence in reversed order. The 252 strings in the primary experimental list were preceded by 18 randomly ordered practice strings. Half the practice strings were words, half were nonwords. None of the practice items were presented more than once.

Procedure

The stimuli were displayed on an Electrohome black-and-white television monitor that was controlled by a Data General computer. Each string was presented inside a small rectangular box that always remained on the screen. The exposure duration for each string was one sec. The interstimulus interval was also one sec. Subjects in the Lexical condition were instructed to press a button marked "yes" if the string presented was a word and to press the button marked "no" if the string presented was not a word. Subjects were told to respond as quickly as possible, but to keep their errors to a minimum (incorrect responses were signalled by a brief flash of the stimulus box). They received no instructions suggesting that there would be any sort of memory test. When asked at the conclusion of the experiment, none of the subjects in the Lexical condition indicated any expectation that there would be receiving a memory test. Subjects in the Memory condition were told to try to remember the information presented on the screen. They received no instructions suggesting that they would receive a frequency test.

Subjects in both the Lexical and Memory conditions received the same frequency discrimination test. A string was presented on the screen, along with the six alternative frequency levels used in the experiment. Subjects were required to select the frequency they thought corresponded to how often the string appeared in the list. They were given all the time they needed for each response before the next string was presented. Half the subjects in each instructional condition received one random order of the 36 test strings (18 words, 18 nonwords). The other half received the same sequence in reversed order.

Results and Discussion

Mean reaction times for subjects in the Lexical group were 559 msec for "yes" responses and 625 msec for "no" responses. The advantage in processing time for "yes" responses was obtained for each subject. It was typical of the Lexical Decision paradigm.

Mean frequency judgments are presented in Table 1. Of primary interest in these data were two results. First, frequency discrimination appeared to be somewhat better for subjects in the Memory condition than for subjects in the Lexical condition. This was indicated by the significant interaction between instructional condition (Lexical vs. Memory) and frequency level (2,4,6,8,10,12), $[F(5,70) = 3.76, \text{ } p < .005, \text{MSe} = 12.66]$. Second, subjects successfully discriminated among the alternative frequency levels following the Lexical Decision orienting task. This was indicated by a test of simple effects, which indicated that the effect of frequency level was significant for subjects in the Lexical condition. Also significant in this experiment was the interaction between string-type (words vs. nonwords) and frequency level, and the three-way

interaction between string-type, frequency level, and instructions
[F(5,70) = 3.10, and [F(5,70) = 2.44, p < .05, MSe = 19.19,
respectively]. These interactions reflect the tendency for frequency
discrimination to be better for nonwords than words, but primarily in
the Lexical condition. No further mention will be made of this effect
of stimulus type since in subsequent research we have continued to
observe a small advantage when subjects receive memory instructions
(compared to completely incidental conditions), but the advantage is
as likely to appear for words as for nonwords.

Finally, correlations between judged and true (list) frequency
were obtained for individual subjects. The correlations were positive
for all 16 subjects participating in this experiment. All eight
correlations were individually significant for subjects in the Memory
condition (they ranged from 0.41 to 0.69). Six of eight correlations
were individually significant for subjects in the Lexical condition
(they ranged from 0.21 to 0.62).

The results of this experiment thus provided further evidence
that frequency information can be accurately encoded under truly
incidental acquisition conditions (the Lexical decision condition).
However, the results also show that nonspecific memory instructions
can elicit strategies, like covert rehearsal, which can enhance
frequency discrimination relative to truly incidental acquisition
conditions.

Table 1

Mean frequency judgments for words and nonwords following
lexical decision instructions or "remember the information
presented" instructions.

Lexical Decision Instructions

Frequency Level

	12	10	8	6	4	2
Words	7.2	7.3	7.8	8.0	4.1	4.2
Nonwords	7.8	6.9	6.9	5.5	5.8	4.8
Mean	7.5	7.1	7.4	6.8	5.0	4.5

Memory Instructions

	12	10	8	6	4	2
Words	8.1	8.1	7.6	5.7	4.8	4.2
Nonwords	9.2	9.1	6.7	6.8	5.3	4.7
Mean	8.7	8.6	7.2	6.3	5.1	4.5

References

Greene, R.L. (1984). Incidental learning of event frequency. Memory & Cognition. 12, 90-95.

Hasher, L. & Zacks, R.T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General. 1979, 108, 356-388.

Hock, H.S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. Journal of Experimental Psychology: Learning. Memory. & Cognition. 12, 232-240.

Zacks have presented interesting data suggesting that frequency encoding can be controlled by variables somewhat different from those of other stimulus attributes; however, they have not demonstrated—by their own criteria—the automaticity of frequency encoding.

Substantial Disruption From Reduction in Capacity

Hasher and Zacks argued that automatic processes should not be disrupted by reduced capacity. Indeed, there is substantial evidence that automatic processes are resource (capacity) insensitive (e.g., see Fisk & Schneider, 1983, 1984; Schneider & Fisk, 1982, 1984). By the "joint satisfaction" rule of Hasher and Zacks, frequency estimation cannot be automatic if reduction in capacity disrupts frequency of occurrence estimations. Hasher and Zacks stated that "reductions in capacity over the ranges so far explored do not affect performance on frequency tests" (p. 1378). Available data clearly contradict this statement.

Fisk and Schneider (1984) carried out an experiment requiring subjects to perform a digit detection visual search task while simultaneously detecting words from a semantic category (e.g., types of vehicles). After substantial practice, untrained exemplars of the trained category were introduced as well as new distractors from other categories. The distractors were presented either 1, 5, 10, or 20 times. In a dual task, subjects were able to detect the untrained exemplars from the trained category with a high degree of accuracy without disrupting the primary digit search task performance. (Those results indicate that subjects were processing the words at least up to the semantic category level in an automatic mode.) However, subjects' estimated frequency of occurrence of the test distractors was independent of the actual presentation frequency. Frequency estimation was relatively good when the subjects' resources were not allocated to the digit search task (Experiment 2 vs. Experiment 1 semantic orienting condition). These data demonstrate that withdrawal of resources from the frequency estimation task disrupts the ability to judge frequency of occurrence. These data do not fit the pattern of results that Hasher and Zacks would need to argue for automaticity of frequency encoding.

Need for Substantial Methodological Care

Fisk and Schneider (1984) have clearly illustrated the requirement for substantial methodological care in attempting to assess relatively pure automatic and non-

automatic (controlled) processes. In light of the publication by Hasher and Zacks, it appears important to reiterate Fisk and Schneider's (1984) requirements. To assess relatively pure automatic processes, researchers must (a) provide evidence as to how well the stimuli actually are processed; (b) provide evidence of the sensitivity of the memory test; (c) provide an excellent cover story; (d) require subjects to perform a highly demanding controlled processing (attention-demanding) task as the primary task; (e) test for frequency estimation only after subjects have learned to allocate attention to the nonautomatic task; and (f) design the task to control for "drift" of attentional resources away from the controlled processing task to the automatic task. The latter three requirements may be met if (a) automatic processes are well developed (e.g., over 2,000 trials of successful executions of the automatized task); (b) subjects are trained to devote full processing capacity to the controlled processing (primary) task; (c) buffer words are presented after automatically processed targets; (d) buffer words are presented at the beginning of each trial to allow time to refocus attention; and (e) highly emotional words (such as rape or murder) are not used.

Summary of Automaticity of Frequency Encoding

Hasher and Zacks (1984) argued that the pattern of null results they reported is critical because their "definition of automatic frequency encoding hinges on the joint satisfaction of six criteria" (p. 1379). (Their criteria predict a pattern of null results.) This comment has pointed to data that indicate that two of their criteria are not supported in the literature. The references cited in this brief comment indicate clear patterns of instruction or strategy effects in the estimation of frequency. Data also clearly show the disruption of frequency encoding when resources are withdrawn from the frequency estimation task. Contrary to the assertion of Hasher and Zacks, it does not appear that frequency information is inevitably encoded into memory.

REFERENCES

Begg, I. (1974). Estimation of word frequency in continuous and discrete tasks. *Journal of Experimental Psychology, 102,* 1046–1052.
Fisk, A. D., & Schneider, W. (1983). Category and word search: Generalizing search principles to complex processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 177–195.
Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 10, 181–197.
Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.
Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372–1388.
Howell, W. C. (1973). Representation of frequency in memory. *Psychological Bulletin, 80,* 44–53.
Postman, L., & Kruesi, E. (1977). The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior, 16,* 353–369.
Rose, R. J., & Rowe, E. J. (1976). Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 142–152.
Schneider, W., & Fisk, A. D. (1982). Concurrent automatic and controlled visual search: Can processing occur without resource cost? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 261–278.
Schneider, W., & Fisk, A. D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 1–15.

Inevitability and Automaticity: A Response to Fisk

Rose T. Zacks
Michigan State University

Lynn Hasher
Temple University

Howard S. Hock
Florida Atlantic University

In the contemporary cognitive literature, the term *automatic* has a number of definitions (e.g., Shiffrin, in press). For the most part, these definitions are not contradictory but complementary to one another. That is, there is significant commonality among them, with the differences arising mainly from the association of the various definitions with research in different cognitive domains. However, mismatches of concerns can occur when different views of automaticity are juxtaposed. Fisk's comment (this issue, pp. 215–216) presents such a mismatch on the topic of methodological problems in the study of automaticity.

As we have used the term automaticity, it refers to a process by which some attributes of an *attended to* stimulus are encoded into memory. We have studied this process mainly through the use of list memory procedures (Hasher & Zacks, 1984, p. 1373). By contrast, the context

of Fisk's comment is a view of automaticity derived from the study of automatic search mechanisms, particularly as they slowly develop in multiple frame visual search tasks (e.g., Fisk & Schneider, 1984). Because of the different foci of the two positions, and especially because of the associated difference in research paradigms, several of the methodological problems that Fisk addresses either do not fit our concerns or are irrelevant to them. The latter is most dramatically illustrated by requirement (a) on page 216: that there be something "over 2,000 trials of successful executions" of a task before it can be assumed that automaticity has been established. In our view, frequency information is encoded into memory (assuming attention to stimuli) on all exposures, including the 1st and the 2,000th.

It is important to note here that we have specified one boundary condition for the obligatory encoding of such fundamental attributes as frequency of occurrence: that the stimuli, although not necessarily the attribute, be attended to (Hasher & Zacks, 1979, p. 359). It may be useful to elaborate on what we mean by the phrase "attended to." We interpret this phrase in a manner consistent with late selection views of attention such as that of Duncan (1980). He argued that stimuli are fully analyzed preattentively, including extraction of their form and meaning. The limited capacity attentional system comes into play to determine which products of preattentive processing will be attended to and thereby brought into consciousness. That is, evidence of some degree of processing of stimuli, even of semantic processing, is not sufficient to demonstrate that the stimuli have been attended to in a manner that meets our boundary condition.

These considerations form the basis for our response to Fisk's claim that reduction in capacity does (contrary to our view) disrupt encoding of frequency information. The data to support this claim come entirely from a study by Fisk and Schneider (1984). In that study it was demonstrated that, given extensive practice, subjects are able to automatically perform even semantically based category searches on words; that is, such searches can be performed on words that are not consciously attended to. If so, Fisk and Schneider's finding that subjects have no memory for the frequency of occurrence of distractors in this paradigm is not contrary to our position: According to our explicit boundary condition, stimuli that are processed in a nonattended way are not expected to leave a record in memory that supports reliable judgments of frequency

of occurrence. In fact, Fisk and Schneider (1984, p. 189) explicitly acknowledged this boundary condition in discussing the relevance of their data to our view of frequency encoding. Thus, it is somewhat surprising that Fisk included this line of argument in his commentary.

We turn now to Fisk's remaining criticism, concerning the impact of instructions on the encoding of frequency of occurrence information. Our instructional criterion states that warning subjects about a forthcoming attributes test will not improve their ability to encode fundamental attributes. This is so because of the presumption that automatic encoding processes function optimally and continuously. Before addressing the issue of whether the data agree with this criterion, we need to clarify a distinction between test instructions and cover task instructions that is honored in the memory literature but is blurred over in Fisk's commentary. There is a difference between instructions about whether to expect a forthcoming memory test (and if so, of what specific type) and instructions about how to process each item as it appears (typically called "orienting" or "cover" tasks). This clarification is necessary to show that the existing data (a) largely conform to our instructional criterion or (b) can be explained by an assumption about subjects' covert rehearsal processes as they proceed through a list of items. We turn first to the impact of test instructions.

Intentional test instructions inform subjects about the nature of the target information that will be tested (e.g., the words themselves, the frequency with which each occurs, their temporal duration or order). Intentional instructions range in the degree to which they go on to specify the actual nature of the forthcoming test from ones that are detailed (e.g., four-alternative, forced-choice item recognition frequency estimation or discrimination, position judgments) to ones that are rather vague, as when subjects are simply told of a "test" without any further information. For our purposes of assessing the impact of instructions on the encoding of fundamental information into memory, intentionally instructed subjects must know that their knowledge of a particular attribute (e.g., frequency) is what is going to be tested.

Incidental instructions are of two types. The first warns subjects only of some unspecified type of test, without any specific information about the target information to be tested. For example, subjects might expect a memory test without knowing that frequency memory will be tested. In the second type of incidental in-

structions, subjects are totally uninformed about a memory test. In this circumstance (sometimes referred to as "truly" incidental) subjects are typically given some task that "orients" them to the items to ensure that the items are actually attended to and that subjects do not guess at the existence of a memory test.

In conformity with the instructional criterion of our framework, encoding of frequency of occurrence information occurs with both intentional and incidental instructions (see Hasher & Zacks, 1984, pp. 1373–1375). Furthermore, recent research of ours shows that encoding of frequency under a number of truly incidental conditions with compelling cover tasks (e.g., a Stroop task, a sentence completion task) is as good as that under incidental and/or intentional instructional conditions (Zacks, Doren, Hamm, Hasher, & Hock, 1985). Two recent articles have also addressed this issue, but they have yielded contradictory conclusions that make their impact unclear. On the one hand, Greene (1984) found that a truly incidental cover task yielded poorer frequency knowledge than the same cover task combined with either vague or explicit memory test instructions. On the other hand, using procedures very similar to Greene's, Kausler, Lichty, and Hakami (1984) obtained a pattern of frequency knowledge in keeping with the automaticity criterion of no instructional differences. Our current conclusion is that, in the main, the results on this variable confirm the automaticity view.

We turn now to the second type of instructional manipulation, which involves varying the type of orienting or cover tasks given to subjects. These instructions are sometimes, though not always, combined with the various types of intentional and/or incidental test instructions, a fact that no doubt contributes to the blurring of the distinction between what are actually two very different sorts of instructional manipulations. A variety of orienting tasks have been used. As examples, subjects may be asked to rate each item as it appears for pleasantness or to indicate the number of syllables each has. Such tasks ensure that subjects pay attention to each stimulus item, but they may also (depending on the particular tasks chosen) result in different amounts of covert rehearsals of the items in the list. For example, subjects who are rating items for pleasantness will try to keep their scale constant across the list and in so doing will rehearse previously presented list items ("Let's see, I think this word is a 6; it's as pleasant as word X, which I also called a 6"). When the cover task directs

attention to individual items, as counting syllables would seem to, fewer rehearsals of prior items occur (see Postman & Kruesi, 1977). Typically (e.g., Fisk & Schneider, 1984, Experiment 1; Rose & Rowe, 1976) the judgments are higher for tasks that encourage rehearsals than for tasks that do not.

We agree that cover tasks differing in the degree to which subjects engage in rehearsals will result in different frequency judgments. We do not see this as a contradiction to our framework because of the following two empirical observations (see e.g., Johnson, Taylor, & Raye, 1977): (a) Subjects are able to judge the frequency of both actual occurrences of items and imagined (or rehearsed) occurrences; and (b) imagined occurrences inflate judgments of actual occurrences (apparently because people sometimes confuse memory traces from the two different sources; see Johnson & Raye, 1981). Thus from our point of view, any variable such as cover task instructions that allows for differential rehearsal rates will set the stage for differential frequency judgments. We have not, as Fisk alleges, ignored this issue, nor have we ignored the relevant data (see Hasher & Zacks, 1984, p. 1380); we seem, however, not to have made ourselves clear. In any event, Fisk's commentary does not attempt to criticize our explanation of the impact of orienting tasks on frequency judgments, and in the absence of such criticism there is no compelling reason to abandon it. His remarks do not present a clear case against our speculations about the special way in which frequency information is encoded.

REFERENCES

Duncan, J. (1980). The locus of interference in the perception of stimuli. *Psychological Review, 87*, 272–300.

Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 181–197.

Greene, R. L. (1984). Incidental learning of event frequency. *Memory & Cognition, 12*, 90–95.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356–388.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39*, 1372–1388.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.

Johnson, M. K., Taylor, T. H., & Raye, C. L. (1977). Fact and fantasy: The effects of internally generated events on the apparent
frequency of externally generated events. *Memory & Cognition, 5*, 116–122.

Kausler, D. H., Lichty, W., & Hakami, M. K. (1984). Frequency judgments for distractor items in a short-term memory task: Instructional variation and adult age differences. *Journal of Verbal Learning and Verbal Behavior, 23*, 660–668.

Postman, L., & Kruesi, E. (1977). The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior, 16*, 353–369.

Rose, R. J., & Rowe, E. J. (1976). Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 142–152.

Shiffrin, R. (in press). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology*. New York: Wiley.

Zacks, R. T., Doren, B., Hamm, V., Hasher, L., & Hock, H. (1985). *The encoding of frequency of occurrence information under truly incidental conditions.* Manuscript in preparation.

Anonymous Reviewing and the Peer-Review Process

Walter W. Surwillo
*University of Louisville
School of Medicine*

It is refreshing indeed to see that the topic of the peer-review process, specifically the matter of anonymous review of manuscripts submitted for publication, is moving from the realm of speculation to the laboratory. Ceci and Peters's comment (December, 1984) is a case in point. They reported an investigation in which reviewers for psychological journals routinely using so-called anonymous reviewing were asked to try to guess the author(s)' identity. Results of this study showed that overall, 35.6% of the 146 participating reviewers were correct in their identification of the author (or one of the authors) of the papers reviewed. These findings were taken as evidence that anonymous reviewing is "fairly blind" and that proponents of anonymous reviews should have confidence in its feasibility.

But does knowledge of authorship ultimately affect publication? Is there a negative bias against unknown authors affiliated with low-prestige institutions and a positive bias in favor of known authors affiliated with high-prestige institutions? These are the critical questions over which the peer-review process has come under attack.

It is regrettable that Ceci and Peters did not carry their study a step further and address this important question. As a

start, it would be nice to know what proportion of the 35.6% of papers whose authors were identified by the reviewers was ultimately published. Is this significantly different from the proportion of the remaining papers (whose authors were not identified correctly) published? It is to be hoped that future studies will wrestle with these questions.

It appears that there are three possible approaches to the peer-review process. The two that have been most frequently employed involve either single-blind review, in which the reviewer knows the identity of the author but the author does not know the identity of the reviewer, and double-blind review, in which neither author nor reviewer knows the other's identity. Because both approaches have elicited so much heated controversy, is it not time to try the remaining alternative, namely, peer review in which the identities of author and reviewer are made known to each other? At the risk of igniting another controversy, I would like to suggest that this may indeed be the fairest and most effective solution to the problem. Why should a reviewer hide behind a cloak of anonymity? If a critique has merit and is really fair, surely the critic ought to have sufficient courage of his or her own convictions to be willing to sign it. Or is anonymous review simply a license for some reviewers to hit below the belt? Certainly open review would be in the tradition of Anglo-Saxon justice, where accuser and accused are able to face each other on equal footing and where openness and the weight of evidence take precedence over rhetoric and reputation.

REFERENCE

Ceci, S. J., & Peters, D. (1984). How blind is blind review? *American Psychologist, 39*, 1491–1494.

The Behavioral Effects of Sugar: A Comment on Buchanan

Richard Milich
University of Kentucky

Scott Lindgren and Mark Wolraich
University of Iowa

Buchanan (November, 1984) labeled refined sugar a "toxin" and called for investigations of the effects of sugar on behavior. He appeared unaware that during the last several years studies have been undertaken to systematically examine the effects of sugar ingestion on the behavior of

SECTION 2F

Emphasis on Global-Level Codes Suppresses the Formation of Component-Level Codes

Howard S. Hock, Lawrence Malcus, and L. Clay Cavedo

Hock, Throckmorton, Webb, and Rosenthal (1981) found that the phonological processing of words (but not of nonwords) suppressed the retention of graphemic information associated with the words. That is, previously presented nonwords that were phonologically processed were recognized more accurately when they were presented in the same case during recognition testing compared with when they were presented in a different case; this advantage of a familiar visual format was not obtained for phonemically processed words. These results suggested that what is remembered in a series of letter strings is influenced by how information in the lexicon is activated during the processing of the strings. When lexical access is based on global-level units (phonological units in the Hock, et al., 1981, study), there is little retention of the visual characteristics of the string; visual characteristics are retained when lexical access is nonphonological, or when there is no lexical activation (for nonwords). Since nonphonological lexical access appears to depend on letter-level visual units (McClelland, 1976; Noice & Hock, 1987), it was hypothesized that experimental conditions which increase the likelihood of direct visual access to the lexicon would facilitate the retention of letter-level memory units, whereas experimental conditions which increase the likelihood of nonvisual (phonological) access to the lexicon would reduce retention of letter-level memory units. Since the hypothesis specifies that lexical activation is critical to differences in letter-level coding, the experimental factors that affect the encoding of letter-level units should be of lesser importance for nonwords, and for tasks which do not require lexical access.

The stimuli presented were randomly mixed lists of words and nonwords. One version of each list was printed entirely in upper-case (e.g., LIST), the other in alternating-case (e.g., LiSt). The reason for introducing the alternating-case condition was to make the stimuli visually unfamiliar, decreasing the likelihood that lexical access would be based on the visual characteristics of the words. For some lists, the nonwords were orthographically regular (e.g., TIBE), for other lists the nonwords were orthographically irregular (e.g., TBEI). The reason for this contrast was that previous research by Shulman, Hornak, and Sanders (1978) suggested that lexical decisions based on a nonvisual, phonological representation of a word were more likely when the nonwords in the list were orthographically regular (and pronounceable) than when they were orthographically irregular (and unpronounceable). There were three cover tasks accompanying the presentation of the words and nonwords: 1) a Letter Detection task in which subjects were required to search through each string for the presence of a "U" or "I", 2) a Lexical Decision task in which subjects were required to discriminate between the words and nonwords, and 3) an Intentional condition in which subjects were told to remember the

frequency of occurrence of the constituent letters of the strings.

Based on the results of a study by Hock, Malcus, and Hasher (1986), the encoding of letter-level memory units was assessed by having subjects estimate the frequency with which various letters appeared in a sequence of strings. However, as pointed out by Hock, et al. (1986), letter frequency estimates could be based on the retrieval of string-level memory units as well as the retrieval of letter-level memory units. That is, using a version of Tversky and Kahneman's (1973) availability heuristic, subjects could base their estimate of a letter's frequency on the number of strings they could recall that included the letter being estimated. Hock, et al. (1986) evaluated this possibility by obtaining free recall data following the frequency estimation phase of their experiment, determining the number of times each letter appeared in correctly recalled strings, and computing reliable partial-correlation coefficients with the effects of recall-frequency "held constant." A similar procedure was used in the present study in order to assess the encoding of letter-level memory units independent of the effects of global-level memory units on the estimation of letter frequency.

Method

Subjects. Two hundred and eighty eight undergraduate students in psychology classes at Florida Atlantic University participated in this experiment, for which they received class credit.

Stimuli. The stimuli presented during the first phase of the experiment were strings of four letters, half of which were words and half of which were nonwords. Sixteen of the consonants composing these strings were designated as target letters; frequency estimates obtained for these consonants during Phase 2 of the experiment provided the data of primary interest for the experiment.

Eight different stimulus lists were constructed. For each list, eight of the target consonants appeared only in words and the other eight appeared only in nonwords. Each set of eight target consonants was further subdivided into four sets of two letters. The average background frequencies of usage (Mayzner & Tresselt, 1965) were similar for the four letter pairs. Each pair of letters was assigned to one of four frequency levels (4, 8, 16, and 32). For example, the letter "P" occurred 16 times by virtue of appearing in 16 different words. The letter "V" also occurred in 16 different words, sometimes in the same word as "P" and sometimes in words with other consonants. Strings containing a target consonant appeared only once in a list. No string contained more than two target consonants, and the same target consonant did not appear more than once in a string. Repetition of a target consonant occurred only after a minimum of two intervening strings that did not include the target. Matching lists were generated by switching the eight target consonants assigned to the words and the eight target consonants assigned to the nonwords (new words and nonwords were generated). The frequency levels to which the target consonants were assigned were also changed in switching the target consonants between words and nonwords.

Additional lists were generated by varying the nature of the nonwords; half were orthographically regular and pronounceable, half were orthographically irregular and unpronounceable. Finally, half

the lists were printed entirely in upper-case, and half were printed
in alternating-case (the format for each string was either
lower-upper-lower-upper or upper-lower-upper-lower, evenly divided
between words and nonwords). As a result of counterbalancing three
factors (the consonants assigned to words or nonwords, the type of
nonwords, the type of case-format), eight different lists were
generated. Thirty six subjects worked with each list, twelve for each
of the three orienting tasks. Subjects were randomly assigned to
forward and reverse versions of each list, with the words and nonwords
randomly ordered within each list.

Each stimulus list was preceded by the same 40 practice strings
(20 words and 20 nonwords in random order). None of the practice
strings included any of the 16 target consonants. Four stimulus lists
contained 160 randomly ordered strings, eight of which were filler
strings which did not include any target letters. The fillers were
used to maintain a minimum of two intervening strings between
occurrences of a target letter. Four stimulus lists had more fillers,
bringing their length to 200 items.

Design. The experiments were conducted in three phases:
acquisition, frequency estimation, and string recall. There were
three Phase 1, acquisition conditions; the second and third phases of
the experiment were identical for all the subjects. The 96 subjects
assigned to the Lexical Decision condition during Phase 1 were
required to determine whether each string was a word or a nonword.
The 96 subjects assigned to the Letter Detection condition during
Phase 1 were required to determine whether or not the letters "U"
and/or "I" were present in each string. The 96 subjects assigned to
the Intentional condition during Phase 1 were told to try and remember
how often each letter appeared in the list. Subjects in the Lexical
Decision and Letter Detection conditions were given no indication that
there would subsequently be any sort of memory test.
Post-experimental interviews confirmed that the memory tests of Phases
2 and 3 were unexpected.

All 26 alphabet letters were presented during the Phase 2
frequency estimation test. Subjects received one of four different
random orders of the letters. All the subjects then participated in
the free recall test of Phase 3.

Procedure. The stimuli were displayed on an Electrohome black
and white monitor that was controlled by a Data General Eclipse
computer. Each string was presented inside a small rectangular box
that always remained on the screen. The box intercepted a visual
angle of 0.7 deg vertically and 2.8 deg horizontally. The exposure
duration for each string was one sec. The interstimulus interval was
also one sec, except for the occasional trials on which subjects
required more than one sec to respond. Then, a one sec delay was
introduced between the subject's response and the presentation of the
next stimulus. Subjects were instructed to press either the button
marked "yes" (for strings that were words in the Lexical Decision
condition; for strings containing a "U" and/or "I" in the Letter
Detection condition) or the button marked "no" (for strings that were
nonwords in the Lexical Decision condition; for strings not containing
a "U" and/or "I" in the Letter Detection condition). In the Lexical
Decision condition, half the strings in both stimulus lists required
"yes" responses. In the Letter Detection condition, 52% of the

strings in four of the stimulus lists and 48% of the strings in the other four stimulus lists required "yes" responses. Half the subjects pressed the "yes" button with a finger of their preferred hand; the other half pressed the "no" button with a finger of their preferred hand. Subjects in these conditions were instructed to respond as quickly as possible, but to keep their errors to a minimum (incorrect responses were signalled by a brief flash of the rectangular box on the screen). There was no overt response required of subjects in the Intentional condition.

During Phase 2, subjects were asked to estimate how often each letter in the alphabet appeared in the preceding list. Subjects presented with one of the four "long" stimulus lists during Phase 1 were instructed to choose a number between 1 and 130 (the most frequent letter, "A", appeared 128 times). Subjects presented with one of the four "short" stimulus list during Phase 1 were instructed to choose a number between 1 and 90 (the most frequent letter, "A", appeared 88 times in this list). The to-be-judged letters appeared one at a time on the Electrohome monitor. Presented alongside the test letter was a number representing the midpoint of the frequency range (65 for the "long" lists, 44 for the "short" lists). Subjects were required to adjust this number upward or downward, using the same buttons as in the Lexical and Letter Detection tasks, in order to estimate the frequency of occurrence for the letter. For subjects presented strings with only upper-case letters during Phase 1, the upper-case version of each letter was presented for estimation during Phase 2. For subjects presented strings with alternating upper-case and lower-case letters during Phase 1, upper-case and lower-case versions of each letter were presented alongside each other for estimation during Phase 2. There was no limit on the time subjects could take for each estimate.

During Phase 3, subjects were provided a blank sheet of paper and given 10 min to recall as many strings as possible from the stimulus list presented during the first phase of the experiment.

Results

The results of the experiment are summarized in Table 1. Since there were no differences among the lists with orthorgaphic and nonorthographic nonwords, the data reported in Table 1 combine the results of those two conditions. As can be seen in the table, estimation accuracy was similar in the Intentional and Lexical Decision conditions, and both were better than estimation accuracy obtained in the Letter Detection condition. The most interesting results were obtained after we "partialled out" the frequency with which the target letters whose frequency was being estimated appeared in subjects' correct-recall protocols. These partial correlation coefficients measured the extent to which actual letter frequency was judged on the basis of letter-level memory units, independent of the contribution of global-level memory units to frequency estimation. We found that partialling-out recall-frequency reduced the size of the correlations in all conditions. This indicated that global-level memory units contributed to the judgment of the frequency of occurrence for the letters in the words and nonwords. However, the most substantial effects of the partial-correlation procedure were

obtained for the alternating-case words in the Lexical Decision
condition. Correlations in this condition were reduced to the point
where they were lower than the correlations obtained in the Letter
Search condition, the reverse of what was obtained prior to
partialling-out the contribution of global-level memory units to
subjects' frequency judgments. This reversal provided evidence for
suppression of letter-level coding in the Lexical Decision condition.

We concluded, therefore, that case-alternation reduced the
likelihood of lexical activation being based on direct visual access,
increasing the contribution of global-level codes (probably
phonological) to the estimation of frequency accuracy, but suppressing
the formation of letter-level memory codes. As expected of an effect
that was based on lexical activation, the suppression of element-level
coding was most evident for words (not nonwords) in a task (Lexical
Decision) requiring lexical activation.

Table 1

Pearson product-moment correlations between actual and estimated
letter frequency.

Phase 1 Orienting Task	Actual with Mean-Estimated Frequency		Mean of Actual with Estimated Frequency for Individual Subjects		Mean of Actual with Estimated Frequency with Recall-Frequency Partialled-Out	
	Words	Non-words	Words	Non-words	Words	Non-words
Intentional						
Same-Case	0.81	0.79	0.51	0.56	0.42	0.55
Alternating-case	0.73	0.64	0.47	0.37	0.35	0.32
Mean	0.77	0.72	0.49	0.47	0.39	0.44
Lexical Decision						
Same-Case	0.75	0.79	0.49	0.48	0.32	0.39
Alternating-Case	0.77	0.62	0.45	0.41	0.17	0.34
Mean	0.76	0.71	0.47	0.45	0.25	0.37
Letter Search						
Same-Case	0.61	0.34	0.41	0.34	0.32	0.19
Alternating-Case	0.68	0.36	0.39	0.31	0.36	0.27
Mean	0.65	0.35	0.40	0.33	0.34	0.23

References

Hock, H.S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. _Journal of Experimental Psychology: Learning, Memory, and Cognition,_ 12, 232-240.

Hock, H.S., Throckmorton, B., Webb, E. & Rosenthal, A. (1981). The effects of graphemic and phonemic processing on the retention of words and nonwords. _Memory & Cognition,_ 9, 461-471.

Mayzner, M.S. & Tresselt, M.E. (1965). Tables of single letter and digram frequency counts for various word-length and letter-position combinations. _Psychonomic Monograph Supplements,_ 1, No. 2, 13-78.

McClelland, J.L. (1976). Preliminary letter identification in the perception of words and nonwords. _Journal of Experimental Psychology: Learning, Memory, and Cognition,_ 2, 80-91.

Noice, H. & Hock, H.S. (1987). A word superiority effect with non-orthographic acronyms: Testing for unitized visual codes. Paper to be read at the annual meeting of the Eastern Psychological Association.

Shulman, H.J, Hornak, R. & Sanders, S. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. _Memory & Cognition,_ 6, 115-123.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. _Cognitive Psychology,_ 5, 207-232.

SECTION 2G

The Encoding of Spatial Relations

Karen Rose and Lynn Hasher

Our research effort has been aimed at the question of whether some
aspects of spatial information, such as spatial relations (e.g., above/
below, along side of) are encoded into memory. Our general strategy has
been to compare memory for relational information across a variety of
tasks such as, frequency judgment, recognition, and priming tasks.
Throughout our work, we found evidence of good memory for the relations
that objects occupied with respect to one another. Further, the ability
to encode spatial relations is not influenced by task variables such as
exposure duration, and cover task. Overall, these findings are important
for at least four reasons: (1) they are the first to systematically
demonstrate that this aspect of spatial information is stored in memory;
(2) they lead to the suggestion that spatial relations play a role in the
encoding of locational information; (3) they show that relational
information is stored in memory under incidental (or unintentional)
conditions; (4) they support other work in our lab in suggesting a view
of space that is different from the Cartesian view of absolute space.

Here follows a brief description of the research which led to our
conclusions.

A. **Frequency Judgments and Spatial Relations**

1. The frequency judgment task was based upon the insight that
peoples' enormous sensitivity to frequency of occurrence information
could be used to determine whether or not a particular unit or element in
the physical world is encoded into memory.

In our first study, we devised a set of 9 context or background
scenes, each containing a relatively large item (e.g., a tv, a crib, a
picnic table). Subjects saw a series of these 9 background scenes in
each of which was placed a single, unique object. The object could occur
in one of several relationships to the context item (including e.g.,
above, below, along side, infront, behind, inside). Across arrays, the
frequency with which a relationship occurred varied (1, 2, and 4). So
for example, a subject may have seen one item above the tv, two items
above the crib, and four above the car. Across the slide series, the
total number of occurrences of each contest was identical; what varied
was the number of times a relationship occurred. Frequency, relationship
and scene were all counterbalanced across subjects.

Subjects were given one of three sets of instructions; relation,
object or incidental. Those given relation instructions were told that
we were interested in peoples' memory for spatial relationships in
scenes. They were then given several examples of the spatial
relationships we were talking about (the particular examples that were
used were matched to those relationships a subject would see). Those in
the object condition were told that we were interested in peoples' memory
for objects in scenes. And those in the incidental condition were told

that they were going to participate in a rating task. Notice that the nature of the test was never mentioned. Immediately after presentation, a forced-choice frequency discrimination test was given; subjects were shown a series of pairs of scenes and were asked to tell in which of the two a particular relation (e.g. above) had been more frequently depicted.

The ability to discriminate the frequency of relationships was reliably above chance for all of the relationships considered (see Table 1). In only one case ws there an effect of instruction. For the relationship behind, subjects in the incidental condition were reliably better in judging frequency fo occurrence, than subjects in the relation condition. Performance for the object instructed group equaled both the relation and incidental conditions. This evidence suggests that relational information is encoded as an incidental byproduct of studying (at least simple) arrays.

2. The first experiment placed a large number of unique items against a small number of backgrounds. In our second attempt to assess subjects' knowledge of relations, a small number of paired objects was used. Each pair occurred in four relations, above/below, side by side, and diagonal to (there were equal numbers of objects placed diagonally right and diagonally left). Further, each pair occurred in each relationship 1, 3 or 6 times. So for example, a subject might see a picture pair (e.g., a bus and a church), one time in an above/below relationship, 3 times in a side by side relationship, and 6 times in a diagonal relationship. Frequency, relationship and pair were all counterbalanced across subjects.

We also varied instructions. One group was told to note the relationship depicted in each pair, but was told nothing about the nature of an upcoming test. A second group was told to note the relationship and was also given full test information. A third group was told only that they would be asked questions about what they had seen. After presentation, all three groups of subjects were asked to estimate how often each pair occurred in each relationship (e.g., how often did bus-church occur in an above/below, side by side or diagonal relationship). Consistent with our prior work, subjects were able to discriminate differences in the frequency with which objects appeared in each relationship. The mean judged frequency for all three relationships increased with actual frequency (see Table 2).

B. Recognition and Spatial Relations

A same/different recognition task was also used to assess memory for relational information. Here, subjects saw simple arrays of two small boxes on a CRT. We explored two issues: (1) sensitivity to three different aspects of spatial information, in particular to exact (Cartesian) location on the screen, to the distance between the boxes, and to spatial relations between the boxes; and (2) the effects of exposure duration (125,250, 500, 1000, or 2000 msec) on performance.

In the experiment, subjects saw an array composed of two small squares set in one of four relationships; above below (:); along side of (• •); diagonal to the left (•.) and diagonal to the right (.•). In addition, each pair could be seen at one of two distances apart, 1 inch and 2 1/4 inches, labeled near and far for convenience.

Subjects received a series of presentation/test trials each of which occurred in the following sequence. First a target pair was presented, followed by a pattern mask, followed immediately by either a foil or the original stimulus. Subjects were required to indicate whether the test item was the same or different for the target item.

To assess spatial knowledge, subjects were given a same/different recognition test for each array immediately after it went off the screen. Across the entire testing series, nine test pairs, eight foils and the target, were constructed for each of the target items. The foils were selected to depict one of three categories of change in spatial information: 1) location changes; 2) proximity changes; 3) relation changes and; 4) the original item itself. Foils that tested for sensitivity to location maintained the distance between squares as well as the original relationship but the array shifted along the x axis (right or left) or along the y axis (up or down). Foils that tested for sensitivity to proximity changes maintained the original relationship but varied the distance between boxes by moving them farther apart (expansion) or closer together (contraction). Foils that tested for sensitivity to relation maintained the distance between boxes but changed their relationship to each other to one of the other three experimental relations. Note that in changing a relation, the exact location of both boxes changed, as they did for proximity and location changes.

The pattern of recognition performance was identical at all exposure durations. For all target types, changes in relationship were detected well above chance and better than either changes in exact location or changes in proximity (see Table 3). Further, proximity changes showed an unusual pattern depending on whether the target items had initially been relatively close together or far apart; subjects were enormously sensitive to contraction movements for near pairs and they were more sensitive to expansion movements for far pairs. In subsequent work, we found that the pattern of performance does not change if the time between target and test item is lengthened (simply by extending exposure of the pattern mask). Nor does it change if a fixation point is provided at the outset of each trial.

3. Priming and Spatial Relations

In our final experiment, subjects studied a series of two-object scenes that were arranged in one of our relationships (above/below, side by side, diagonal to and front/behind) (see Appendix 1). They were instructed to study each object pair carefully, because after presentation, memory for object information would be tested. After presentation of the study sequence, subjects were given a recognition test which was made up of the original scenes (whole scenes), part of the

original scene (one subject), parts of new scenes (one object), or new scenes comprised of two new objects (whole scenes). The subject's task was to decide whether the object or objects had appeared in the study list. Response and latency to respond were recorded.

Two aspects of the test procedure varied. The first was whether the target was in the same or a different relationship as that shown at study. Targets appeared in the same relationship as that shown at presentation, or in one of the three other experimental relationships. Note however, the exact location of all target items changed at test by shifting the pair to the right or left. In this way, we hoped to look at relational information, independent of exact location. The second aspect that varied at test was whether the target item was primed or misprimed. When the target was to be primed, a part of the scene (one subject) directly preceded the target. When the target was misprimed, an item that was presented during presentation, but was part of some other filler scene preceded the target. Along with never-presented items, in which part of an actually presented scene was followed by a pair which had never been presented, there was a total of five trial types: primed-same relationship; primed-different relationship; misprimed-same relationship; misprimed-different relationship; and negative control items.

For three of the four presentation relations, above/below, side by side, diagonal to, relationship at test was a significant factor. In all cases, reaction time was faster when the relationship was maintained for presentation and test, in comparison to when it changed. The relationship at test factor was not significant for front/behind however, In addition, for all four input relations, there was a main effect of trial type. In all cases the pattern was the same; reaction time was faste ron those trials in which a part of the scene directly preceded the target, than on those in which an unrelated scene preceded the target. In sum then, time to decide whether a target item had appeared int he study list depended on whether the spatial relationship ahd been maintained from presentation to test (even though subject shad only to remember the object sin order to do the task), and on what item had preceded the target.

Taken together, the results of our studies suggest that relational information may well be an important aspect of what is encoded about objects in an array.

Table 1

Proportion Correct on Forced Choice Test of Knowledge of Frequency as a Function of Instructional Condition

	Object	Relation	Incidental
Relationship at Test			
Above	.62	.59	.65
Below	.64	.64	.69
Side by Side	.62	.65	.63
Infront[a]	.67	.61	.66
Behind[a]	.62	.60	.71*
Inside	.70	.70	.71

[a] The relationships infront and behind are collapsed across variations of each. That is, one object could be placed straight infront of another, infront but diagonally left of another, or infront but diagonally right of another. Similarly, one object could be placed straight behind another, behind but diagonally left of another, or behind but diagonally left of another. The significant effect of instructions was actually produced for behind diagonally right.

* $p < .05$

Table 2

Mean Judged Frequency as a Function of Actual Frequency
and Relationship Collapsed Across Instructions

	Actual Frequency		
Relationship	1	3	6
Side by Side	2.76	3.58	4.07
Above/Below	3.07	3.76	4.35
Diagonal To	2.69	3.26	3.74

Table 3

Percent Error as a Function of Relationship, Foil Type, and
Initial Location (Near or Far) Collapsed Across Presentation Rate

Relationship*	Location		Proximity		Relationships			
	up/down	left/right	exp#	cont#	SS	AB	DL	DR
Near								
SS	.26	.42	.19	.11	--	.07	.13	.07
AB	.32	.40	.30	.11	.09	--	.07	.07
DL	.33	.36	.26	.11	.05	.06	--	.09
DR	.32	.35	.23	.06	.07	.05	.13	--
MEAN	.31	.38	.25	.10	.07	.06	.11	.08
Far								
SS	.25	.29	.25	.48	--	.03	.01	.04
AB	.24	.24	.19	.41	.06	--	.03	.04
DL	.24	.32	.14	.28	.01	.04	--	.07
DR	.31	.41	.26	.21	.05	.04	.05	--
MEAN	.26	.32	.21	.34	.04	.04	.03	.05

* SS = side by side
  AB = above below
  DL = diagonal left
  DR = diagonal right

# exp = expansion
  cont = contraction

Table 4

Mean Latency to Respond and Proportion of Errors (in parentheses) as a Function of Relationship at Presentation, Relationship at Test, and Trial Type

| Relationship at Presentation	SR*	DR
	Primed Trials	
ABOVE/BELOW	1274.49 (.10)	1300.04 (.11)
SIDE BY SIDE	1289.81 (.02)	1333.39 (.07)
DIAGONAL TO	1278.57 (.09)	1313.78 (.08)
FRONT/BEHIND	1268.88 (.13)	1293.18 (.08)
	Misprimed Trials	
ABOVE/BELOW	1347.00 (.05)	1394.66 (.13)
SIDE BY SIDE	1371.57 (.07)	1450.50 (.13)
DIAGONAL TO	1345.45 (.06)	1402.48 (.06)
FRONT/BEHIND	1361.29 (.10)	1366.71 (.06)
	Negative Trials	
ABOVE/BELOW	1716.00 (.06)	
DIAGONAL TO	1774.71 (.08)	
FRONT/BEHIND	1811.31 (.06)	

* SR = SAME RELATIONSHIP
  DR = DIFFERENT RELATIONSHIP

Note: Proportion of errors reflect incorrect responses only. This includes responding "no" in the case of a primed or misprimed trial and "yes" in the case of a negative trials.
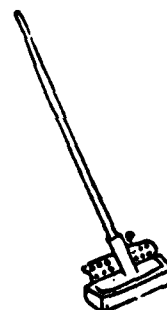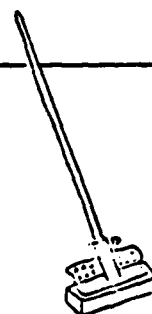
Appendix

Sample Stimulus Materials. From left to right:
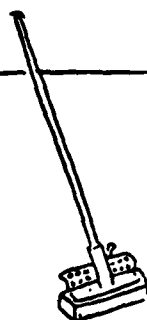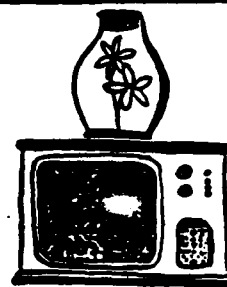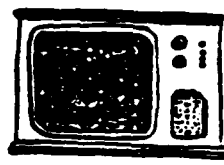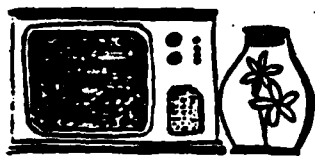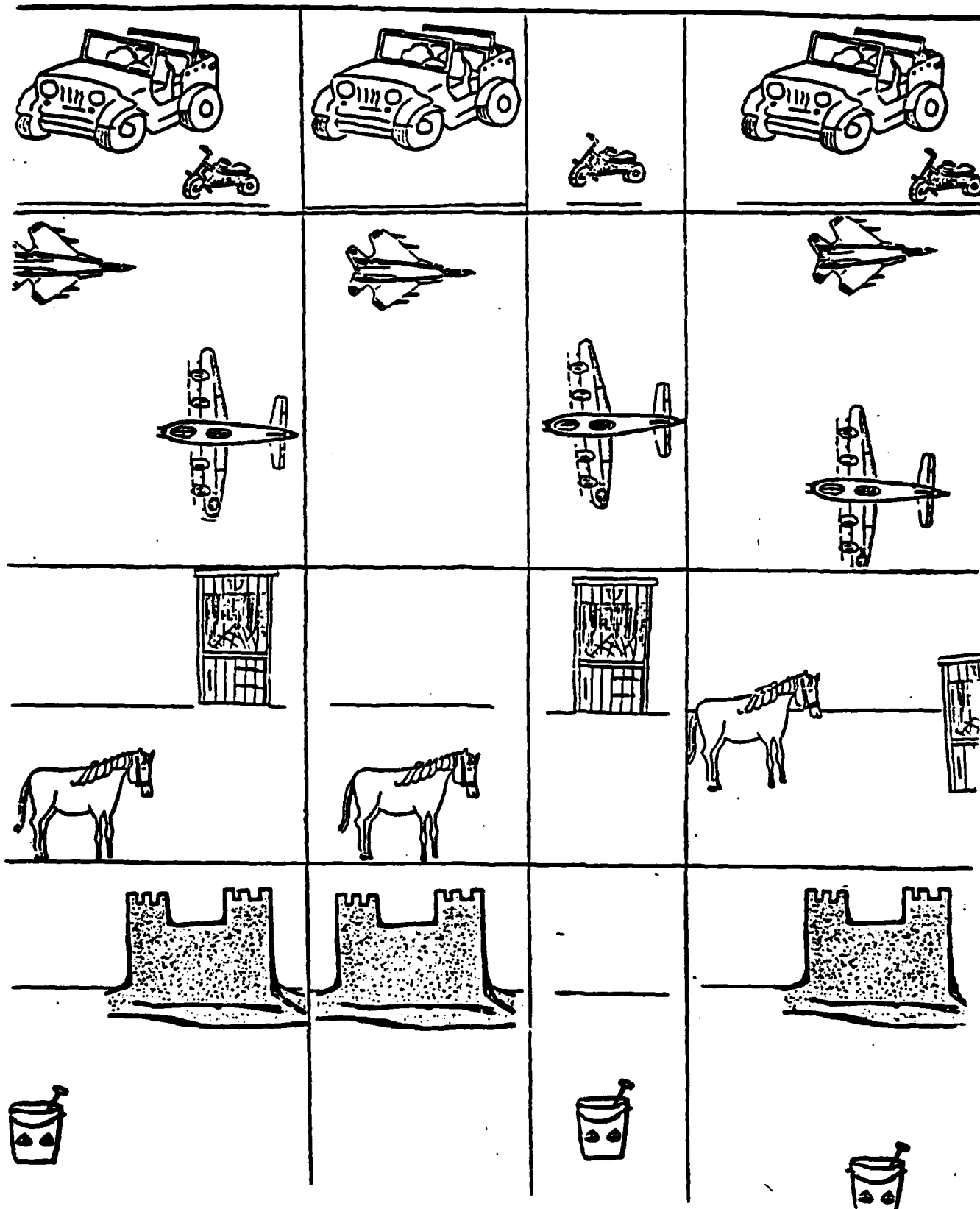Presentation Pair (Column 1); Primes (Columns 2 and 3);
and Test Pair (Column 4). Note: exact location
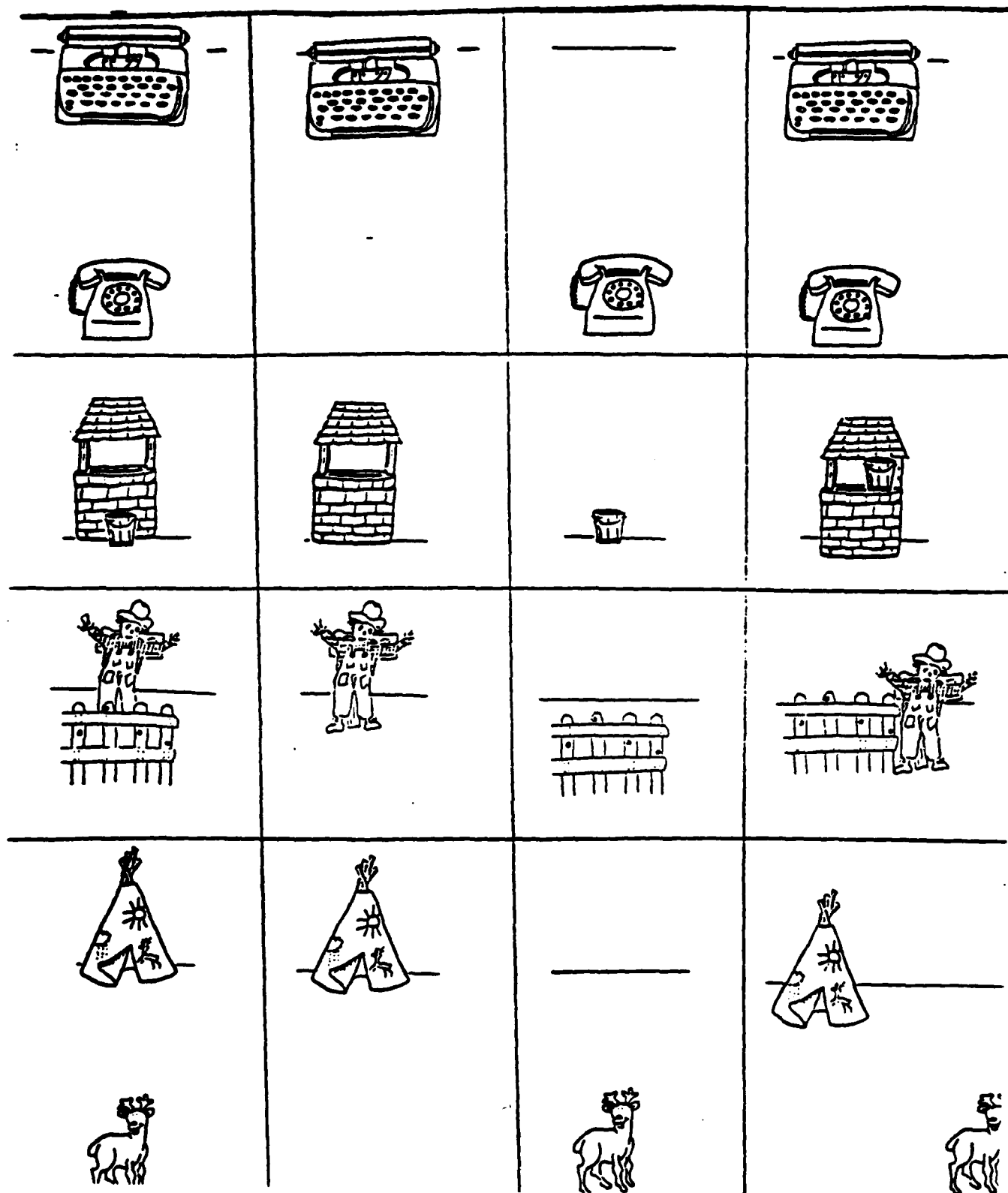changed in all cases.

Presentation Pair: ABOVE/BELOW
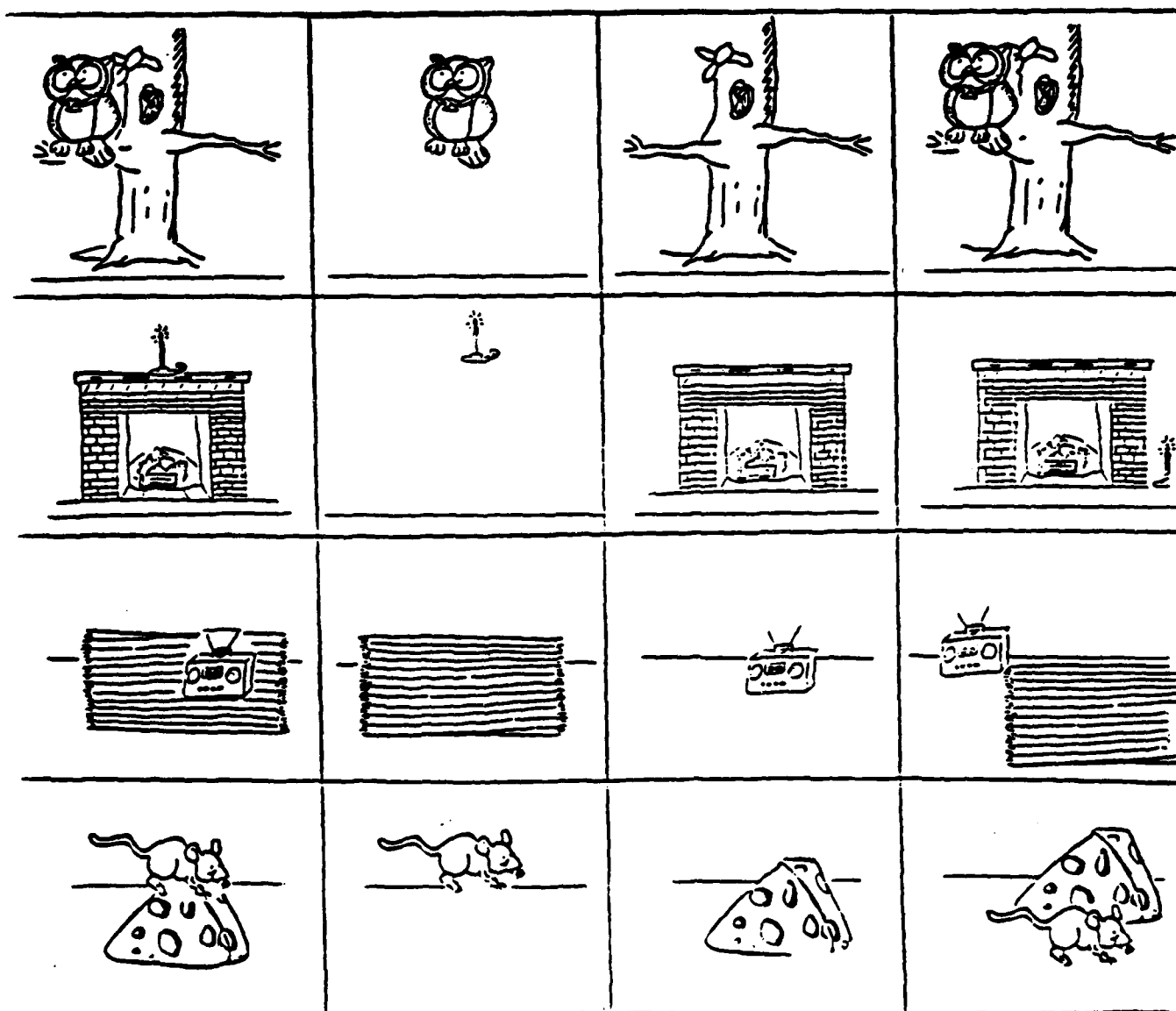
Presentation Pair: SIDE BY SIDE

Presentation Pair: FRONT/BEHIND

SECTION 2H

The Effect of Pattern Structure on Frequency Judgments

Howard S. Hock, Alexandra Bates, and Linda Field

Patterns with simple, symmetrical structures are easier to identify (Clement, 1964) and reproduce (Bell & Handel, 1976) than patterns with more complex, asymmetrical stuctures. In this experiment we examined whether similar differences would be observed when subjects judged the frequency of occurrence of simple and complex patterns, and the extent to which the effect would depend on subjects' orienting task and the location of the patterns within a surrounding frame.

Method

Sixteen different patterns, each composed of five circles (within an imaginary 3 x 3 matrix), were presented within a 4x4 frame with 16 possible locations for the circles. These were the same patterns used in the Hock, Smith, Escoffery, and Bates (Section 2C) study. Half the patterns were relatively simple, the simplicity resulting from the symmetry of the patterns [in terms of Garner & Clement's (1964) analysis, these patterns belonged to an equivalence class of size 4]. Half the patterns were relatively complex; they lacked symmetry [in terms of Garner & Clement's (1964) analysis, these patterns belonged to an equivalence class of size 8].

For each subject, eight patterns (4 simple; 4 complex) were presented four times, and eight patterns (4 simple; 4 complex) were presented eight times. Each subject therefore saw a random sequence of 96 patterns. Eight of the patterns (half simple, half complex; half presented eight times, half presented four times) were always presented in one of the four quadrants of the 4x4 frame; eight of the patterns (half simple, half complex; half presented eight times, half presented four times) were equally distributed among the four quadrants of the frame.

Subjects participated in one of two tasks during the initial presentation of the patterns. One group searched for the presence of a dot within one of the circles, the other was told to try to remember the patterns. This was followed by a frequency estimation test during which patterns were presented without the surrounding frame.

Results

As can be seen in Table 1, subjects did a better job of discriminating between the low and high frequency patterns in the Pattern Memory compared with the Dot Detection condition. Whether or not the patterns appeared in the same or different quadrants did not affect frequency estimation. Although subjects discriminated between high and low frequency "simple" patterns to a greater extent than they discriminated between high and low frequency "complex" patterns, the interaction between frequency level and complexity fell short of statistical significance at the .05 level.

Table 1
Mean estimates for frequency of occurrence of patterns varying in complexity.

| | SIMPLE PATTERNS | | | | COMPLEX PATTERNS | | | |
| | Same Quadrant | | Different Quadrant | | Same Quadrant | | Different Quadrant | |
	Low Freq (4)	High Freq (8)	Low Freq (4)	High Freq (8)	Low Freq (4)	High Freq (8)	Low Freq (4)	High Freq (8)
PATTERN MEMORY	5.0	6.8	5.0	6.2	5.0	5.7	4.4	5.4
DOT DETECTION	4.2	4.7	4.5	5.0	5.3	5.5	5.5	5.9

References

Bell, H.H., & Handel, S. (1976). The role of pattern goodness in the reproduction of backward masked patterns. Journal of Experimental Psychology: Human Perception and Performance. 2, 139-150.

Clement, D.E. (1964). Uncertainty and latency of verbal naming responses as correlates of pattern goodness. Journal of Verbal Learning and Verbal Behavior. 3, 150-157.

Garner, W.R. & Clement, D.E. (1964). Goodness of pattern and pattern uncertainty. Journal of Verbal Learning and Verbal Behavior. 2, 446-452.

Hock, H.S., Smith, L., Escoffery, L., and Bates, A. (1985). Pattern perception: Coding the position of component elements. Paper read at the 26th annual meeting of the Psychonomic Society, Boston.

SECTION 3A

A Word Superiority Effect With Non-orthographic Acronyms:
Testing for Unitized Visual Codes

Helga Noice and Howard S. Hock

Abstract

Letters in briefly presented, masked letter-strings were detected more accurately when the strings were three-consonant acronyms than when they were nonwords. In the absence of orthographic regularity, this word superiority effect (WSE) could not have depended on visual units corresponding to familiar bigrams. However, evidence that the WSE was obtained only for letters in the third position indicated that it did not depend on the formation of whole-word visual units. It was argued instead that the WSE for acronyms depended on the post-lexical activation of sequential, associatively connected, single-letter phonological codes. Finally, the results of case-alternation, size-alternation, and mixed type-font experiments were interpreted in conjunction with the view that lexical access is based on both lower-case and upper-case letter recognition units for words, and only upper-case letter recognition units for acronyms.

In Reicher's (1969) tachistoscopic recognition paradigm, the brief presentation of a test string (e.g., WORK) is preceded and followed by masking characters that interfere with the processing of the string. Subjects are then provided with two response alternatives (e.g., WORK-WORD). Their choice between the alternatives, which in the above example would indicate whether they detected the presence of a "K" in the last letter of "WORK," is better for strings which are words than strings which are nonwords. This result, which has been termed the word superiority effect (WSE), indicates that tachistoscopic letter-detection is facilitated by the activation of lexical entries for the test word.

Carr and Pollatsek (1985) have recently proposed that the WSE results from the formation of nonvisual, whole-word unitizing codes that "protect" the information in the briefly presented word from the effects of masking and memory loss. Evidence consistent with this hypothesis has been reported by Hawkins, Reicher, Rogers, and Peterson (1976), who showed that letter-detection was significantly better than chance for standard response alternatives (e.g., WORD-WORK), but was at chance level for phonologically identical response alternatives (e.g., SITE-CITE). Another possibility is that the WSE could result from the formation of unitized, whole-word codes that are visual rather than phonological. Experiments in which consecutive letters in a test string are presented in different cases (e.g., wOrK) are consistent with the hypothesis that higher-order visual codes could mediate the WSE. Pollatsek, Well, and Schindler (1975) found that case-alternation increased the time required for subjects to detect letter differences between pairs of words, but not between pairs of orthographically irregular nonwords. Although these results showed that case-alternation can affect lexical activation by rendering words visually unfamiliar, McClelland (1976) found that the advantage of words compared to orthographically regular pseudowords was not affected by case-alternation. This indicated that if visual unitization was responsible for the WSE, the units were smaller than the whole word. It is possible, however, that whole-word visual unitization was not observed in McClelland's study because lexical activation based on whole-word visual codes was less efficient than lexical activation based on phonological codes (formed by the application of spelling-to-sound translation rules to the orthographically regular words).

The purpose of the experiment reported in this paper was to determine whether a WSE could be obtained for the rapid, masked presentation of three-consonant acronyms (e.g., NBC). The use of these acronyms provided a strong test of the "visual unitization hypothesis" for two reasons. First, their shortness and the consistency with which they have been experienced in purely upper-case format increased the likelihood of the acronyms being recognized (i.e., their lexical entries activated) on the basis of whole-word visual units. Second, the absence of vowels made it impossible for lexical activation to depend on whole-unit phonological codes formed via spelling-to-sound translation rules. If evidence for the existence of whole-word visual codes cannot be obtained for these stimuli, it is unlikely that such evidence would be obtainable for orthographically regular words. The latter are usually longer than acronyms, the likelihood of recognizing them on the basis of whole-unit visual codes is decreased because they are experienced in different case formats (sometimes all upper-case, sometimes all

lower-case, sometimes with the first letter upper-case and the remaining letters lower-case), and they lend themselves to the formation of unitized phonological codes via spelling-to-sound translation rules. If lexical access depends on the relative efficiency of parallel phonological and visual processes (see Carr & Pollatsek, 1985, for a review of parallel coding systems models), there may be too many factors favoring the phonological channel for orthographically regular words.

Previous experiments have provided evidence for a WSE with acronyms (Egeth & Blecker, 1971; Henderson, 1974), but the effect has been unreliable (Carr, Posner, Pollatsek, & Snyder, 1979), subject to response bias (Henderson & Chard, 1976; Seymour & Jack, 1978), and dependent on whether the acronyms are presented to the left of right visual field (Besner, Davelaar, Alcott, & Parry, 1984). None of these experiments used the Reicher (1969) paradigm, and most of their items were at least partially orthographically regular (e.g., US in USA). We eliminated vowel-based orthographic combinations in our experiment in order to reduce the likelihood of partial spelling-to-sound translations, and because the presence of visually familiar bigrams might reduce the likelihood of subjects activating lexical entries for the acronyms on the basis of whole-word visual units. The inconsistency of the WSE for previous studies using acronyms might have been due to subjects forming intermediate-level visual codes corresponding to orthographically familiar letter combinations within the acronyms (Glushko, 1979). The WSE would not be reliably obtained if lexical entries for the acronyms cannot be activated through these intermediate-level codes.

If we found superior letter-detection performance for the acronyms compared with three-consonant control strings, it would indicate that the WSE can be obtained in the absence of any orthographic regularity. Further evidence that the size of the WSE could be reduced by rendering the acronyms visually unfamiliar (via case-alternation) would be consistent with the hypothesis that the WSE for acronyms depends on lexical activation via whole-word visual codes. However, the critical test for the whole-word, visual unitization hypothesis would be to obtain the WSE, and the effects of case-alternation on the WSE, regardless of whether letter-detection is tested in the first or third position of the acronyms.

Method

Subjects

The subjects were 36 Florida Atlantic University psychology students who participated in the experiment as one alternative to fulfilling a course requirement. Each subject spoke English fluently and had normal or corrected-to-normal vision.

Stimuli

The 10 acronyms used in this experiment were all composed of three consonants (see Table 1). A preliminary study indicated that subjects recognized each of them (they were the 10 most familiar among a set of 30 three-consonant acronyms that were examined).

The first step in constructing the nonword, control strings was to group the acronyms into pairs. For each pair, the first two letters of one acronym were combined with the third letter of the other acronym, or the last two letters of one acronym were combined with the first letter of the other acronym. Thus, pairing NBC with LSD yielded NBD and LBC as the nonword-controls for NBC, and NSD and LSC as the nonword-controls for LSD. The 10 acronyms were paired-off in this manner to generate 20 nonword-controls (see Table 2). Using the same procedure, four more all-consonant acronyms (GNP, FTC, KGB, and CPR) were used to generate eight nonword-controls for the practice list. For one group of subjects, all the letters were presented in upper-case. For a second group of subjects, all the strings were presented in alternating-case format (lower-upper-lower).

Design

The probe letter for each trial appeared equally often in the first and third positions. Half of the trials for each position were "yes" trials and half were "no" trials. Each acronym was tested in both the first and third positions. For NBC, the probe letter for the first position was N for "yes" trials and L for "no" trials; for the third position, the probe letter was C for "yes" trials and D for "no" trials. Each nonword-control string was tested in either the first or third positions, but not both. As can be seen in Table 2, when the probe letter for testing a nonword-control string did not correspond with the letter being tested (a "no" trial), the probe letter in combination with the other two letters in the string formed one of the acronyms from which the nonword-control was generated (e.g., the third-position "no" probe for NBD was C, which in combination with NB formed NBC). This feature of the design was the basis for testing familiarity-bias in subjects' responses (further discussion of this test is presented in the Results section).

As can be seen in Table 2, there were two different nonword-control strings for each acronym, but the total number of trials was the same for the acronyms (each was tested in two positions) and the nonword-controls (each was tested in only one position). Since each position was tested twice, once with a "yes" probe and once with a "no" probe, each of the 10 acronyms was tested four times and each of the 20 nonword-controls was tested twice, producing a total of 80 stimuli. Three random orders of these stimuli produced three blocks of 80 trials. Each subject worked on one of three orders (Latin square) of the three blocks. These 240 experimental trials were preceded by up to 96 practice stimuli, these constituting three blocks of 32 trials formed from the four practice acronyms and their eight nonword-controls.

Procedure

Testing took place in a semi-darkened, partially sound-proof room. An Apple IIe microcomputer was used to present the stimuli and record

subjects' responses. A shield placed on the computer keyboard exposed two response keys (labelled "yes" and "no") and the "return" key. Each three-letter string was 0.5 cm in height and 1.2 cm in width. When viewed from a distance of 1.3 m, each string subtended a visual angle of 0.2 deg vertically and 0.5 deg horizontally.

At the beginning of the experimental session, the experimenter told the subjects that they would be seeing three-letter strings, some of which would be meaningful and some of which would be meaningless. They were also told that a single letter would be presented immediately after each string, and that they were to decide whether or not the letter had appeared in the string in the same position as the test letter. The sequence of each trial was as follows: a fixation point in the center of the screen was replaced by three ampersands (&&&), which were presented for 850 msec. The ampersands, in turn, were replaced by a three-letter string. After the string was presented (see below for the duration), the three ampersands replaced the string, along with a single probe letter beneath either the first or third position of the string. This display remained on the screen until the subject responded.

The experiment began with the presentation of up to 96 practice stimuli. Subjects responded by pressing one of two keys (feedback was provided by a "beep" after an incorrect response). The practice trials involved a psychophysical staircase procedure, the purpose of which was to select an exposure duration for each subject that would result in his/her detection accuracy reaching an asymptote of approximately 70% correct. The staircase procedure resulted in the selection of an exposure duration for each subject that was used for all 240 experimental trials. The mean exposure duration was 88 msec for strings presented entirely in upper-case and 98 msec for strings presented in alternating-case. The difference between the two conditions was not significant, [$t(34) < 1.0$].

- - - - - - - - - - - - - - -
Insert Table 3 About Here
- - - - - - - - - - - - - - -

Results

Mean percentage errors and d' scores for the experimental trials are presented in Table 3. In the analyses of variance that follow, Fs refers to tests against error terms based on subject variability and Fi refer to tests against error terms based on item variability.

A WSE was obtained, but only when letter-detection was tested in the third position of the three-letter strings. The results based on percentage errors indicated that there was a significant interaction between stimulus-type and position [$Fs(1,34) = 30.01$, $p < .001$, MSe = 11.82; $Fi(1,9) = 46.89$, $p < .001$, MSe = 15.43]. Similar results were obtained when the analysis was based on d' scores [$Fs(1,34) = 20.04$, $p < .001$, MSe = .41; $Fi(1,9) = 44.85$, $p < .001$, MSe = .12]. Tests of simple effects (on percentage errors) indicated that the acronym/nonword difference (i.e., the WSE) was significant for the detection of letters in Position 3 [$Fs(1,34) = 27.56$, $p < .001$, MSe = 13.33; $Fi(1,9) = 42.55$,

p < .001, MSe = 15.54], but was not significant for the detection of letters in Position 1 [Fs(1,34) < 1.0; Fi(1,9) < 1.0].

Case-alternation reduced the size of the WSE by a factor of two, the reduction being particularly evident when letter-detection was tested in the third position. However, the effect of case-alternation on the WSE was only marginally reliable. The interaction between stimulus-type and case-type on percentage errors was significant when tested against subject variability [Fs(1,34) = 5.46, p < .05, MSe = 29.66], but fell just short of significance when tested against item variability [Fi(1,9) = 4.27, p > .05, MSe = 34.13]. When the analysis of variance was based on d' scores, the interaction fell just short of significance when "items" was the random factor [Fi(1,9) = 4.56, p > .05, MSe = .44], and was not significant when "subjects" was the random factor [Fs(1,34) = 2.69, p > 05, MSe = .34]. The marginal reliability of the case-alternation effect on the WSE could have been the result of case-alternation having its strongest effects when letter-detection was tested for the first position, but the WSE was obtained only when letter-detection was tested in the third position.[1]

- - - - - - - - - - - - - -

Insert Table 4 About Here

- - - - - - - - - - - - - -

The consistency of the WSE in Position 3 is highlighted by the d' data in Table 4; a positive WSE was obtained for each of the 10 acronyms. However, three of the items, FDR, NFL, and FPL, might be construed as having some residual orthographic regularity; the last two letters of each are orthographically regular bigrams. Their bigram regularity notwithstanding, we considered their occurrence irregular because DR, FL, and PL never appear in the final positions of words. However, to give our hypothesis the strongest possible test, we performed an additional analysis without these three items (and their associated nonword-controls). Although there was an over-all reduction in accuracy, the pattern of results and the outcome of statistical analyses remained the same.

The final analysis tested for familiarity-bias. Henderson and Chard (1976) and Seymour and Jack (1979) found that subjects were biased to make "same" as opposed to "different" responses when the strings they were comparing were familiar acronyms compared with unfamiliar control strings. Carr, Posner, Pollatsek, and Snyder (1979) obtained a similar familiarity-bias for orthographically regular words. For this reason, we tested for whether subjects were biased to respond "yes" when the probe letter combined with the previously detected information in the string to form a familiar acronym. Such a bias could have resulted in letter-detection being more accurate for the acronyms than the control strings.

Our procedure for determining whether the WSE was due to familiarity-bias was based on detection performance for the nonwords. Consider the string LBC. The probe letter for testing detection of the first letter in this string was either L or N. The latter, combined with detected information specifying that there was a B and C in the second and third positions of the string, could produce the familiar string,

NBC. A higher rate of false alarms ("yes" responses to N) than misses
("no" responses to L) would indicate that subjects were biased to respond
"yes" when the probe letter completed a familiar acronym. However, this
conclusion could be reached only in the absence of a general bias to make
more "yes" than "no" responses. In testing our nonword data for
familiarity-bias, we found that the false alarm rate was equal to or
lower than the miss rate at both probe positions and for both case-types.
In addition, there was no indication of a general bias toward "yes"
responses (only 48.9% of the responses were "yes"). It was unlikely,
theefore, that the WSE we obtained was the result of familiarity-bias.

Discussion

The experimental results indicated that a WSE could be obtained with
three-consonant acronyms. The absence of vowels for these stimuli made
it impossible for the WSE to have depended on whole-unit phonological
codes formed on the basis of spelling-to-sound translation rules.
Furthermore, since the WSE did not require the presence of
orthographically regular spelling combinations, it could be concluded
that if it depended on visual unitization, the units formed could not
have corresponded to intermediate-level, orthographically familiar
bigrams (Glushko, 1979). Nonentheless, the data did not support the
hypothesis that the WSE for the acronyms was mediated by whole-word
visual units. The WSE was obtained only when letter-detection was tested
in the third position. If the WSE for the acronyms depended on the
formation of whole-word visual units, it would have been obtained for all
letter positions that were tested. Evidence that letter-position
affected detection accuracy for the nonwords (implying left-right
scanning), but not the acronyms, might be construed as supportive of the
hypothesis that whole-word visual units were formed for the acronyms.
However, case-alternation reduced the size of the WSE without introducing
the left-right scanning effects suggested by the nonword data. If the
WSE for the acronyms depended on the formation of whole-word visual
units, and case-alternation was interfering with the formation of such
units, then left-right scanning effects would also have been observed for
the case-alternated acronyms.

In the absence of evidence for the formation of whole-word visual
units, and the impossibility of forming intermediate-level visual units,
we concluded that the WSE obtained for the acronyms was mediated by the
post-lexical activation of higher-order phonological units. The units
proposed would not be formed as a result of spelling-to-sound translation
rules; the acronyms lacked the vowels necessary for the application of
such rules. Instead, they would involve the names for each letter in the
acronym, with phonological unitization resulting from the activation of
sequential, associatively connected, single-letter phonological codes
stored in the lexical representation for the acronym.

One of the reasons Carr and Pollatsek (1985) proposed their
unitization hypothesis was the need to protect the coded information in
the briefly displayed strings from memory loss until the post-stimulus
response alternatives could be evaluated. If sequential, associative
connections among letter names were serving this function, it would be
expected that the facilitative effects of familiarity would be observed

for letters near or at the end of the string; these would be the letters most susceptible to being forgotten in a left-right scan of single-letter phonological codes. This was the case in the experiment reported in this paper, as well as an unreported, replication experiment (see again Footnote 1); a WSE was obtained for the acronyms, but only for the detection of letters in the third position.

The effect of case-alternation in reducing the size of the WSE indicated that access to the lexicon was visually mediated. However, the absence of orthographic regularity and the failure to obtain the WSE in both test positions indicated that lexical access was based on letter-level visual units. But why does case-alternation reduce the WSE for acronyms (also reported by Besner et al., 1984), and not for orthographically regular words (McClelland, 1976)? As indicated earlier, orthographically regular words are experienced in a variety of ways, sometimes entirely lower-case, sometimes entirely upper-case, and sometimes with the first letter upper-case and the remaining letters lower-case. As a result, it is proposed that the lexical representations for orthographically regular words include recognition units for both the upper- and lower-case versions of their constituent letters. It would be for this reason that case-alternation has no effect on the WSE for orthographically regular words. Because acronyms are experienced exclusively in upper-case, the lexical entries for acronyms would include recognition units for only the upper-case versions of their constituent letters. It would be because of the introduction of lower-case letters that case-alternation reduces the WSE for acronyms (it may not eliminate it because lexical activation could still occur via letter-level phonological codes). Alternating the size of consecutive upper-case letters would not affect the WSE for acronyms (Besner, et al., 1984) if the attributes for recognizing the upper-case letters were specified with sufficient abstraction for the recognition units to generalize to letters of any size. A similar argument would account for the invariance of the WSE when words are presented in mixed type-fonts (Adams, 1979).

In conclusion, the results of this experiment indicated that a WSE could be obtained in the absence of orthographic regularity. However, there was no support for the hypothesis that the WSE for acronyms was mediated by whole-word visual codes. As argued earlier, if evidence for whole-word visual unitization could not be obtained for three-consonant acronyms, it is unlikely that such evidence would be obtainable for orthographically regular words. Our evidence against the formation of whole-word visual units is therefore at odds with one of the fundamental assumptions underlying the "sight" method of reading instruction (see Crowder, 1982, for review). Instead, of visual unitization, our experimental results were explained in terms of phonological unitization. This was consistent with Carr and Pollatsek's (1985) assertion that unitization in the WSE is typically phonological. In the case of acronyms, phonological unitization appears to be based on the post-lexical generation of sequential, associatively connected phonological codes for individual letters, rather than the application of spelling-to-sound translation rules to orthographically regular letter strings. The lexical representations containing these associatively unitized phonological codes seem to be accessed by letter-level visual units, but they might also be accessible via letter-level phonological units.

References

Adams, M. J. (1979). Models of word recognition. <u>Cognitive Psychology</u>, <u>11</u>, 133-176.

Besner, D., Davelaar, E., Alcott, D., & Parry, P. (1984). Wholistic reading of alphabetic print: Evidence from the FDM and the FBI. In L. Henderson (Ed.), <u>Orthographies and reading: Perspectives from cognitive psychology, neuropsychology and linguistics</u>. Hillsdale, N.J.: Erlbaum Associates.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, and G. E. MacKinnon (Eds.), <u>Reading research: Advances in theory and practice</u>, <u>5</u>, Orlando, Florida: Harcourt, Brace, Jovanovich.

Carr, T. H., Posner, M. I., Pollatsek, A., & Snyder, C. R. R. (1979). Orthographic and familiarity effects in word processing. <u>Journal of Experimental Psychology: General</u>, <u>108</u>, 389-414.

Crowder, R. G. (1982). <u>The psychology of reading: An introduction</u>. New York: Oxford University Press.

Egeth, H., & Blecker, D. (1971). Differential effects of familiarity on judgments of sameness and difference. <u>Perception & Psychophysics</u>, <u>9</u>, 321-326.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>5</u>, 674-691.

Hawkins, H. L., Reicher, G. M., Rogers, M., & Peterson, L. (1976). Flexible coding in word recognition. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>2</u>, 380-385.

Henderson, L. (1974). A word superiority effect without orthographic assistance. <u>Quarterly Journal of Experimental Psychology</u>, <u>26</u>, 301-311.

Henderson, L., & Chard, J. C. (1976). On the nature of the facilitation of visual comparisons by lexical membership. <u>Bulletin of the Psychonomic Society</u>, <u>7</u>, 432-434.

Pollatsek, A., Well, A. D., & Schindler, R. M. (1975). Familiarity affects visual processing of words. <u>Journal of Experimental Psychology: Human Perception & Performance</u>, <u>1</u>, 328-338.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. <u>Journal of Experimental Psychology</u>, <u>81</u>, 275-280.

Seymour, P. H., & Jack, M. V. (1978). Effects of visual familiarity on "same" and "different" decision processes. <u>Quarterly Journal of Experimental Psychology</u>, <u>30</u>, 455-469.

Author Notes

Footnotes

[1]The results reported in this experiment were replicated as one part of another experiment. In this replication, which included two more acronyms (BLT, VHF) and their associated nonword-controls, the WSE was again obtained only when letter-detection was tested in the third position. Analyses of variance on percentage errors replicated the significant interaction between stimulus-type and position [$F_s(1,34)$ = 28.76, $p < .001$, MSe = 33.06; $F_i(1,11)$ = 12.41, $p < .05$, MSe = 110.59]. Tests of simple effects indicated that the acronym/nonword difference in letter-detection accuracy was significant for Position 3 [$F_s(1,34)$ = 45.33, $p < .001$, MSe = 37.32; $F_i(1,11)$ = 19.83, p .01, MSe = 120.25], but not for Position 1 [$F_s(1,34)$ = 2.34, $p > .05$, MSe = 37.23; $F_i(1,11)$ = 1.10, $p > .05$, MSe = 120.25].

Case-alternation again reduced the size of the WSE, and the reduction was again greater when letter-detection was tested in the third position. However, in this experiment the effect of case-alternation on the WSE was statistically reliable: the interaction between stimulus-type and case-type on percentage errors was significant [$F_s(1,34)$ = 15.10, $p < .001$, MSe = 41.39; $F_i(1,11)$ = 8.92, $p < .05.$, MSe = 100.26].

Table 1

Acronyms used for the experimental trials.

LSD	Lysergic Acid Diethylamide
NBC	National Broadcasting Company
LBJ	Lyndon Baines Johnson
FDR	Franklin Delano Roosevelt
JFK	John Fitzgerald Kennedy
PBS	Public Broadcasting System
CBS	Columbia Broadcasting System
FPL	Florida Power and Light
NFL	National Football League
BMW	Bavarian Motor Works

Table 2

Items and probe letters for the experimental trials.

		TEST LETTER			
ITEMS		YES RESPONSE		NO RESPONSE	
Acronyms	Nonwords	Pos. 1	Pos. 3	Pos. 1	Pos. 3
LSD		L	D	N	C
	NSD	N		L	
	LSC		C		D
NBC		N	C	L	D
	LBC	L		N	
	NBD		D		C
LBJ		L	J	F	R
	FBJ	F		L	
	LBR		R		J
FDR		F	R	L	J
	LDR	L		F	
	FDJ		J		R
JFK		J	K	P	S
	PFK	P		J	
	JFS		S		K
PBS		P	S	J	K
	JBS	J		P	
	PBK		K		S
CBS		C	S	F	L
	FBS	F		C	
	CBL		L		S
FPL		F	L	C	S
	CPL	C		F	
	FPS		S		L
NFL		N	L	B	W
	BFL	8		N	
	NFW		W		L
BMW		B	W	N	L
	NMW	N		8	
	BML		L		W

Table 3

Mean percentage error rates and d' scores.

Stimulus Type	Upper-Case		Mixed-Case	
	1st Position	3rd Position	1st Position	3rd Position
PERCENTAGE ERRORS				
Acronyms	8.1	10.4	16.7	13.6
Controls	10.2	24.1	14.9	21.2
Difference (WSE)	2.1	13.7	-1.8	7.6
d'				
Acronyms	2.91	2.82	2.25	2.51
Controls	2.84	1.60	2.30	1.81
Difference (WSE)	.07	1.22	-.05	.70

Table 4

The word superiority effect (WSE) for each acronym obtained
by calculating the difference in d' between each acronym
and its nonword control strings.

	Upper-Case		Alternating-Case	
	1st Position	3rd Position	1st Position	3rd Position
LSD	1.79	2.93	-0.98	0.07
NBC	0.52	1.53	-0.80	0.88
LBJ	0.51	0.89	-0.31	0.85
FDR	0.00	2.24	0.87	1.32
JFK	-0.23	1.29	0.94	0.94
PBS	0.32	1.60	1.61	0.74
CBS	0.44	1.98	-0.76	0.71
FPL	0.66	1.70	-0.67	2.05
NFL	-0.03	1.34	0.70	0.53
BMW	0.13	0.63	-1.22	0.08

SECTION 3B

Perceptual Units in the Acquisition of Visual Categories
Howard S. Hock, Cheryl Tromley, and Lynn Polmann

Abstract

A series of experiments provided evidence that the representational
structure of categories comprising dot patterns is based on pattern-parts
and pattern-configuration rather than pattern-elements. We found that
similarity judgments and post-acquisition classification data could not
be explained in terms of element-level perceptual units, even for
categories of dot patterns with seven of their eight dots in the exact
same relative location. The importance of higher-order perceptual units
was indicated by evidence that the long-term retention of information
specific to previously learned category exemplars, which is typical of
natural objects, can also be obtained for artificial dot patterns,
providing their structure reflects the perceptual characteristics
identified in Tversky and Hemenway's (1984) study of natural objects:
members of the same category had to be perceptually distinctive at the
level of pattern-configuration, and perceptually similar at the level of
pattern-parts. The level of within-category similarity for a set of
categories (relative to between-category similarity) did not predict
whether item-specific information would be retained; long-term retention
appears to require both within-category similarity and dissimilarity, but
at different levels of perceptual structure.

Contemporary research concerned with categorization has had two major foci. The first, which can be traced to the research of Rosch and her colleagues (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), has been concerned with the attribute structure of categories comprising natural, real-world objects. The second focus, which received its impetus from the work of Posner and Keele (1968, 1970), has been concerned with the acquisition of categories comprising initially unfamiliar, artificial stimuli (typically dot patterns). Although there have been studies that have bridged these two approaches (e.g., Rosch, Simpson, & Miller, 1976; Murphy & Smith, 1982), our understanding of category structure for natural and artificial categories has been divergent with respect to the long-term retention of information specific to previously learned category exemplars, and with regard to what is considered to be the appropriate level of perceptual analysis for determining the representational structure of the categories.

Our everyday experience with natural objects is sufficient to demonstrate that being able to categorize an object is compatible with being able to distinguish it from other members of its category. For example, being able to classify an object as a dog is compatible with being able to identify particular, well-known dogs and being able to determine that an object, though readily classifiable as a dog, has never been seen before. In contrast, the retention of information specific to individual category exemplars has not been evident in research involving the acquisition of artificial categories. The typical experimental result is that previously learned category exemplars enjoy an advantage relative to novel category members when classification testing occurs immediately after acquisition, but the advantage is reduced or eliminated when classification testing is delayed for relatively long periods of time (typically, one week). This pattern of classification data has been observed for such diverse stimuli as dot patterns (e.g., Posner & Keele, 1970), random polygons (Homa, Sterling, & Trepel, 1981), photographs of real-world scenes (Hock & Schmelzkopf, 1980), and painting style (Hartley & Homa, 1981).

The research reported in this paper is based on the view that the long-term retention of item-specific exemplar information depends on perceptual relationships among the objects belonging to the categories. Tversky and Hemenway's (1984) work with natural objects suggests that critical perceptual relationships occur at the level of an objects' constituent parts. They've argued that objects belonging to the same basic level category can differ in overall shape or configuration, but their category membership depends on shared attributes determined at the level of the objects' parts. In contrast, researchers working with artificial stimuli have generally argued that the category membership of a stimulus depends on its similarity to a central, prototypical representation (e.g., Posner & Keele, 1968). Implicit in this view, which has been formalized in conjunction with a model proposed by Knapp and Anderson (1984), is the assumption that the location of individual elements is the appropriate level of analysis for artificial dot patterns. Knapp and Anderson (1984) propose that categories are learned via the cumulative neuronal effect of the sensory inputs associated with a particular response. When applied to dot patterns, their model assumes that each dot produces neuronal activity that is maximized at the perceived location of the dot and diminishes

exponentially in surrounding locations. Patterns of dots are "represented as spatially organized bumps of activity (p. 623)" that are added to the neuronal activity produced by previous visual inputs associated with the same response. Therefore, when the elements of successive patterns remain in similar locations, well-defined peaks of neuronal activity emerge at those locations, but when the relative spatial locations of the dots are distorted in successive patterns, the neuronal activity associated with each dot is maximized at a location between the various locations of the dots. In this way, Knapp and Anderson (1984) account for the emergence of prototypical information in memory without the assumption of separately stored exemplars and prototypes. Although they acknowledge that pairwise and higher-order relations among the elements of a dot pattern could contribute to the pattern's memory representation, and their model, at least on an abstract level, could accomodate higher-order units, its most concrete application is for element-level units.

In the present study, we were concerned with whether the appropriate unit of perceptual analysis for studying categorization is at the level of individual elements, which has been implicit in most studies based on artificial dot patterns, or at the level of parts, as indicated by studies of natural object categories. Our initial objective was to demonstrate, contrary to Knapp and Anderson's (1984) formalization, that perceptual relations among dot patterns belonging to the same category are not adequately represented at the level of individual elements. Our second objective was to show that the long-term retention of information specific to previously learned category exemplars, which is typical of natural objects, can also be observed for artificial dot patterns. The necessary condition for the latter would be that the categories encompassing the artificial patterns be constructed to reflect the higher-order perceptual units that are the basis for the categorization of natural objects. Based on Tversky and Hemenway's (1984) description of natural object categories, it was hypothesized that evidence for the long-term retention of information specific to individual category exemplars would be obtained if dot patterns belonging to the same category were similar to each other at the level of the patterns' parts, and uniquely different from each other the level of the patterns' configuration. Our determination of configural similarity and uniqueness, was based on the correspondence of pattern-part That is, two patterns were considered configurally similar if they were composed of the same parts, and the parts were in the same relative location. They were considered configurally distinct when they were composed of unique combinations of parts.

Experiment 1

In this experiment we describe the three sets of patterns used in the experiments reported in this paper, the way subjects segmented the patterns into their constituent parts, and finally, how subjects judged the relative similarity of patterns for each of the three sets. These results were the basis for Experiments 2 and 3, which examined acquisition and post-acquisition classification performance for each of the three sets of categories, the objective being to test for the long-term retention of information specific to individual category members.

Stimuli

Dot patterns were defined as belonging to the same category if they were generated from the same base pattern. The four base patterns used to generate the categories are presented at the top of Figure 1. Each base pattern comprised eight dots whose locations were defined by positions in an imaginary 11 x 11 grid. The spacing between grid lines was equal to one dot-diameter, resulting in patterns that were relatively compact. For descriptive convenience, the categories generated from the four base patterns were identified by a letter (A, B, C, or D), and each member of a category was identified by a number (1 through 8). Three sets of categories (I, II, and III) were generated.

Set I. The patterns belonging to the four categories in Set I are presented in Figure 1. The procedure used to generate these patterns is illustrated in Figure 2. Category members were generated by selecting one dot from each base pattern and moving only that dot to new locations outside its immediate neighborhood in the base pattern. Eight different patterns were generated by placing the "movable" dot in eight different locations. As a result, for each of the patterns belonging to the same category, seven of the eight dots were in precisely the same relative location. The locations for the "movable" dot were selected to create categories for which individual patterns had distinctive configurations, despite their high level of 1:1 element-correspondence.

- - - - - - - - - - - - - - - - - -
Insert Figures 1 and 2 about here
- - - - - - - - - - - - - - - - - -

Set II. The patterns in Set II, which were generated from the same base patterns as Set I, are presented in Figure 3. Category members were produced by selecting a group of three dots from the base patterns that were located near each other, and moving these dots one space in the same direction (see Figure 4). For each base, two category members were generated by moving three dots to the left, two were generated by moving three dots to the right, two by moving three dots up, and two by moving three dots down. A different grouping of three dots was selected for movement to produce each of the eight category members (each dot in each base pattern was moved equally often).

- - - - - - - - - - - - - - - - - -
Insert Figures 3 and 4 about here
- - - - - - - - - - - - - - - - - -

Set III. The Set III patterns, which are presented in Figure 5, were generated by a more severe version of the Set II procedure. The first step was to segregate the eight dots composing each base into two groups of three dots and one group of two dots. The dots within each group were then moved one space in the same direction, but the different groups were each moved in a different direction (see Figure 6). Each pattern generated in this way involved a different segregation of the base pattern's eight dots into groups of three, three, and two, and a different pattern of movement-direction for the three groupings.

Analysis of Pattern-Parts

Method. Three groups of 128 undergraduate students at Florida Atlantic University participated in this part of the experiment. They were required to segment dot patterns into their constituent parts. One group was assigned to the Set I patterns, one group to the Set II patterns, and one group to the Set III patterns. Each subject worked with only four patterns, one from each category (A, B, C, or D). They were not shown more than one member of a category because we were concerned that they might recognize a pattern as a new version of a preceding one and feel constrained to repeat their preceding responses. Booklets comprising four 14 cm x 20 cm sheets of white paper were formed. A different dot pattern (2.5 cm x 2.5 cm) was presented on each sheet. The four patterns in each booklet always came from different categories in the following eight combinations: A1B1C1D1, A2B2C2D2, A3B3C3D3, A4B4C4D4, A5B5C5D5, A6B6C6D6, A7B7C7D7, A8B8C8D8. Sixteen booklets were formed for each combination (the order of patterns within each booklet was counterbalanced), resulting in the production of 128 booklets. Subjects were unaware that the patterns were to be used in categorization experiments. Some were tested individually; most were tested in groups.

Each subject was given a booklet of four patterns and instructed to draw circles around the dots in each pattern to form three clusters (the number of clusters was restricted so that each subject's data would have an equal weight in determining the perceptual parts of each pattern). Subjects were told that the clusters should reflect their perception of the natural groupings of dots within the patterns. The clusters they formed could overlap, and one cluster could be contained within another. Every dot had to appear in at least one cluster, even if it appeared alone, and no circle could surround all the dots in the pattern.

Results. Our analysis of the parts of each pattern focussed on the three clusters of elements that were circled most frequently by subjects (the three clusters encompassed all eight dots composing each pattern). It was assumed that they were the most salient parts or subunits of the pattern. For all three sets, we found that patterns belonging to the same category shared the same salient parts to a much greater extent than patterns belonging to different categories.[1] Closer examination indicated that the part-structures for the Set I and Set III categories were similar. For both sets, pairs of patterns from the same category generally shared one or two salient parts, but different pattern-pairs usually shared different salient parts. Despite this part-level over-lap, there was only one instance in which the three most salient parts of a pattern corresponded to the three most salient parts of another pattern. Thus, every pattern shared some of its parts with other members of its category, but almost every pattern was composed of a different combination of parts. We concluded on this basis that members of the same Set I and Set III categories were configurally unique.

In contrast with the Set I and Set III patterns, we found that the three most salient pattern-parts appeared in the same combination (and the same relative-location) in the majority of Set II patterns belonging

to the same category. On this basis, we concluded that most members of the same Set II categories were configurally similar.

Similarity Judgments

Method. Six undergraduate and graduate students at Florida Atlantic University participated in this experiment. None had foreknowledge that the patterns were to be used in categorization experiments. Each subject rated perceived similarity for a randomly ordered sequence of 360 pattern-pairs; 120 pairs from Sets I, II, and III. The 120 pairs from each set were formed from all combinations, both within- and between-category, of Patterns 1, 2, 7, and 8. The assignment of each pattern to the left and right position of each pair was balanced across the 120 pairs, each of which was presented on a 14 cm x 20 cm sheet of white paper.

Procedure. Subjects' judgments were based on a 10-point rating scale, 1 indicating that the patterns were highly dissimilar and 10 indicating that they were highly similar. Subjects were given several minutes to look through the stack of pattern-pairs before they began.

- - - - - - - - - - - - - -
Insert Table 1 about here
- - - - - - - - - - - - - -

Results. As indicated in Table 1, mean similarity ratings for within-category pattern-pairs from Set I were intermediate to the ratings obtained for Set II and Set III. The same ordering was independently observed for each of the six subjects. There was little difference in between-category similarity ratings among the three sets.

Discussion

Although seven of the eight dots were in exactly the same relative location for patterns belonging to the same Set I category, subjects rated these pattern-pairs as lower in similarity than within-category pattern-pairs from the same Set II category. This suggested that similarity ratings for pairs of Set I patterns did not depend on the correspondence of element-level perceptual units. To argue otherwise, it would have to be assumed that similarity judgments were more affected by large differences in location for one dot (the Set I patterns) compared with small differences in location for many dots (the Set II patterns). However, an inspection of the Set I patterns in Figure 1 makes it clear that the detection of the one discrepant ("moving") dot is not easy, even when one is looking for it. Since subjects judging similarity had no reason to look for the discrepant dot, it is unlikely that their judgments were biased to emphasize the relatively large differences in its location. Similarity ratings for these patterns were, in all likelihood, influenced by higher-order perceptual units than individual pattern-elements.

The analysis of pattern-parts indicated that the structure of the Set I categories reflected the perceptual characteristics of natural object categories that we hypothesized are necessary for the long-term retention of information specific to individual category members. That is, patterns belonging to the same category resembled each other at the

part-level, and were perceptually distinctive at the configural-level. These perceptual requirements were not well established for Sets II and III. For the Set II categories, patterns belonging to the same category resembled each other at the part-level, but in most cases were not as perceptually distinctive at the configural-level. For the Set III categories, patterns belonging to the same category were perceptually distinctive at the configural-level, but judgments of within-category similarity for these patterns indicated that their part-level resemblance were relatively weak compared with the Set I and Set II categories (the arrangement of dots within each part and the location of the part within its pattern was more likely to be exactly the same for pairs of patterns belonging to the same Set I and Set II categories compared with pairs of patterns belonging to the same Set III categories).

On the basis of these results, it was predicted that evidence for the long-term retention of information specific to individual category exemplars would be obtained for the Set I categories, but the more typical experimental result, a loss of item-specific information with delayed testing, would be obtained for the Set II and Set III categories. These predictions were tested for the Set I and Set II categories in Experiments 2a and 2b, and the Set I and Set III categories in Experiments 3a and 3b. The methodology used in these experiments was originated by Posner and Keele (1970). It involved introducing a variable delay between category acquisition and classification testing, and contrasting classification performance for previously learned (original) and novel category members, the latter being generated by the same rules as the previously learned category members. A significant difference in classification accuracy between the original and novel members of the same category would indicate that subjects retained information specific to the previously learned (original) patterns.

Since Medin, Dewey, and Murphy (1983) have shown that the nature of the training procedure can affect original/novel differences in classification accuracy, two acquisition procedures were used. One required learning the same name for each category member, the other different names for each category member.

<div align="center">Experiment 2a</div>

Method

Subjects. Fifty-six undergraduate students at Florida Atlantic University voluntarily participated in this experiment.

Design. The patterns belonging to each Set I category were divided into two subsets, (1-4) and (5-8). Subjects were randomly assigned to one of eight conditions defined by the orthogonal combination of training procedure (concept-formation vs. paired-associate), pattern subset during acquisition (1-4 vs. 5-8), and time of classification testing (immediate vs. two-week delay). Classification testing involved the presentation of 36 patterns: 16 original patterns that were presented during acquisition, 16 novel category members that had not been seen before, and 4 base patterns that also had not been seen before. If subset (1-4) comprised the originals during acquisition, subset (5-8) comprised the novels during

testing, and vice versa. The 36 patterns were presented in one of two random sequences. Matching sequences were provided for counterbalancing the (1-4) and (5-8) subsets as originals and novels.

Procedure. Black on white slides of each pattern were back-projected, via a random-access projector, onto a translucent screen (their size was 2.5 cm x 2.5 cm). Viewed from a distance of approximately 114 cm, they intercepted a visual angle of 1.3 deg.

Two types of acquisition procedure were used. In concept-formation training, subjects learned the same verbal label for each member of a category (see Table 2). They were told that they would be seeing 16 dot patterns, each of which belonged to one of four groups, and were instructed to look for relations among the patterns that would allow them to learn which patterns belonged in each group. In paired-associate training, subjects learned different verbal labels for each pattern (see Table 2), but the names for members of the same visual category belonged to the same semantic category. The labels for Category A were all names of colors, Category B names of cities, Category C names for money, and Category D names of months. Subjects receiving paired-associate training were not instructed to look for visual relations among the patterns (although the semantic similarity of the pattern names could have directed attention to visual relations among members of the same category).

- - - - - - - - - - - - - -
Insert Table 2 About Here
- - - - - - - - - - - - - -

Acquisition began with subjects being shown 16 patterns in random order (four patterns from each category), the experimenter orally identifying each pattern. Subsequent blocks of 16 trials were presented in one of five random sequences, the order of these sequences varying from subject to subject. Ten seconds were given for subjects to orally identify each pattern, at which point they were required to make their best guess. Corrective feedback was provided following incorrect responses. The acquisition phase continued until two consecutive blocks of 16 trials were completed without error.

Classification testing followed acquisition immediately, or after a two-week delay. Each pattern was presented for as long a period of time as subjects needed to classify the pattern. They were required to guess if they were unsure of the correct response. No corrective feedback was provided. If subjects had received concept-formation training during acquisition, they were told that they would be seeing some old patterns and some new ones they hadn't seen before, and that the patterns belonged in the same four groups that they had already learned. They were provided with the appropriate verbal labels (red, green, blue, and yellow) on an index card. If subjects had received paired-associate training during acquisition, they were told that they would be seeing some old patterns and some new ones that they hadn't seen before, and that the patterns belonged in four distinct groups: color, city, money, and month (the appropriate labels were again provided on an index card). By using the names of the semantic categories, both the previously seen and novel category members could be identified with the same responses.

Results

Subjects receiving concept-formation and paired-associate training required an average of 6.5 and 10.4 trial-blocks (excluding the two errorless criterion blocks), respectively, to reach the acquisition criterion (all subjects reached criterion).

- - - - - - - - - - - - - -
Insert Table 3 About Here
- - - - - - - - - - - - -

As can be seen in Table 3, classification performance was more accurate after paired-associate training than after concept-formation training. This difference was probably the result of the longer training period required by the paired-associate condition. Following both training procedures, original category members (from the acquisition set) were classified more accurately than previously unseen, novel category members, the original/novel difference being maintained over the two-week delay.[2] An analysis of variance indicated that whether the patterns were original or novel significantly affected percent classification accuracy [$F(1,52) = 20.35$, $p < .001$, $MSe = 83.98$], but did not significantly interact with the time-of-test (immediately vs. delay), [$F(1,52) < 1.0$, $MSe = 83.98$]. The advantage of the original compared with the novel category members was independently observed for each of the four Set I categories (A, B, C, and D). The effect of training procedure (concept-formation vs. paired-associate) was significant [$F(1,52) = 5.06$, $p < .05$, $MSe = 121.48$], but none of the interactions with training procedure were significant. Finally, classification accuracy was somewhat lower for the base patterns than other category members that weren't seen before. There was no indication that the base patterns had any special status in this experiment compared with other category members that were not seen before.

Discussion

The results of this experiment indicated that previously learned, Set I patterns were classified more accurately than novel, Set I patterns, and moreover, that this difference was unaffected by the introduction of a two-week delay between acquisition and classification testing. This evidence for the long-term retention of item-specific information was obtained regardless of whether learning to identify patterns individually was an explicit requirement of the training task (paired-associate training) or not (concept-formation training). The experiment therefore supported the hypothesis that long-term retention of exemplar information could be obtained for artificial dot patterns when patterns belonging to the same category were perceptually similar to each other at the part-level and perceptually distinct at the configural-level.

Experiment 2b

Method

This experiment involved acquisition and classification testing for the Set II patterns. With the exception that the Set II patterns were substituted for the Set I patterns, the design and procedure were identical

to that of Experiment 2a. Fifty-six undergraduate students at Florida Atlantic University voluntarily participated in this experiment.

Results

Subjects receiving concept-formation and paired-associate training required an average of 2.5 and 17.8 trial-blocks (excluding the two errorless criterion blocks), respectively, to reach the acquisition criterion (all subjects reached criterion). Concept-formation training was easy because members of the same category, which were to receive the same name, were so similar. However, it was because they were so similar that it was difficult to learn distinctive names for each category member, as required by paired-associate training.

- - - - - - - - - - - - - - -
Insert Table 4 about here
- - - - - - - - - - - - - - -

An examination of the classification errors, which are presented in Table 4, indicated that performance was again more accurate after paired-associate training than after concept-formation training. In addition, the effect of time-of-test (immediate vs. delay) on the classification of originals and novels was comparable following concept-formation and paired-associate training. Although the effect was more clear-cut following concept-formation training, for both training procedures a difference in classification performance between original and novel category members was obtained for immediate testing, but not when testing took place two weeks after acquisition.

The classification data obtained in this experiment did not lend themselves to analysis of variance. This was because a large number of subjects (68% in the immediate condition, 43% in the delay condition) did not make any errors for either the original or novel patterns. Instead of analysis of variance, we collapsed the data over training condition, determined whether each subject made more errors on novels or originals, and performed a sign test that contrasted the novel-original differences for immediate and delayed testing. We found a significantly greater tendency toward a positive novel-original difference in immediate compared with delayed testing [$\chi^2(2) = 6.78$, $p < .05$]. In the immediate condition, nine subjects made errors on novel and/or original patterns; eight made more errors on novels than originals, and one was tied. A sign test indicated that the probability was only .04 that eight of nine subjects would have made more errors on novels than originals as a result of chance variation. Furthermore, the higher error rate for novels compared with originals was consistently obtained within each of the four Set II categories (A, B, C, and D). In the delay condition, 16 subjects made errors on novel and/or original patterns; seven made more errors on novels, six made more errors on originals, and three were tied. A sign test indicated that the probability was .80 that seven of 16 subjects would make more errors on novels than originals as a result of chance variation. Thus, a significant difference between the classification of novel and original patterns was obtained when testing occurred immediately after acquisition, but the difference was not significant with delayed testing. Finally, ceiling effects made it difficult to evaluate the data obtained

for the base patterns. The results obtained when testing occurred immediately after concept-formation training suggested that information specific to the previously learned category members (originals) generalized to the base patterns to a greater extent than they generalized to the novel patterns.

Discussion

The reduction in the original/novel difference in classification accuracy with delayed testing was consistent with previous literature using the Posner and Keele (1970) paradigm. The results supported the hypothesis that the lack of configural distinctiveness for the Set II categories would impede the long-term retention of information specific to previously learned patterns. In addition, obtaining the typical original/novel by time-of-test interaction for the Set II categories showed that the atypical results obtained for the Set I categories in Experiment 2a were not an artifact of our experimental procedure.

Experiment 3a

In this experiment we again used the Set I patterns, but strengthened the test of long-term retention by extending the duration between acquisition and classification testing to four weeks. Another difference in procedure compared with Experiment 2a was that instead of presenting the base patterns during classification testing, we presented patterns which were formed by averaging the locations of dots from patterns belonging to the same category (these were called central patterns). If subjects' representations of the Set I categories were based on the locations of individual elements (recall that seven of the eight dots in each category member were in the exact same relative location), classification accuracy would be better for the central patterns than the novel category members, which also were never seen prior to the testing phase. The third change in procedure involved the introduction of yes/no recognition testing in combination with classification testing. This provided an additional measure of whether subjects retained information specific to the patterns presented during acquisition. Finally, only concept-formation training was used in this experiment (the procedure differed from Experiments 2a and 2b for reasons unrelated to the presented study).

Method

All 36 subjects, again undergraduate students at Florida Atlantic University, received concept-formation training with the Set I categories. One group learned Categories A and B first, then Categories C and D (each phase had a criterion of two consecutive blocks of eight errorless trials). A second group first learned C and D, then A and B. This was followed, for both groups, by a third acquisition phase in which 16 patterns from all four categories were presented in randomly mixed sequences until subjects completed two consecutive blocks of 16 errorless trials.

The subjects were further subdivided into three groups, one of which was tested immediately after acquisition, one after a two-week delay, and one after a four-week delay. During testing, subjects were required to classify each pattern and then indicate whether or not they had seen it

during the preceding acquisition phase (due to experimenter error, only 7 of 12 subjects in the immediate testing condition received recognition testing; all 12 received classification testing). As before, subjects were required to guess if they were unsure of the correct response, and no corrective feedback was provided.

- - - - - - - - - - - - -
Insert Table 5 about here
- - - - - - - - - - - - -

Results

A total of 8.5 trial-blocks were required for category acquisition: 2.8, 2.3, and 3.4 for the first, second, and third training phases, respectively (again excluding errorless criterion blocks). As before, all subjects successfully reached the acquisition criterion.

The classification and recognition data are presented in Table 5. As in Experiment 2a, there was no apparent effect of delayed testing on the original/novel difference in classification accuracy. An analysis of variance indicated that the original/novel difference [$F(1,33) = 18.93$, $p < .001$, MSe $= 66.01$] and the effect of time-of-test [$F(2,33) = 4.55$, $p < .05$, MSe $= 374.19$] were significant, but the original/novel by time-of-test interaction was not significant [$F(2,33) < 1.0$, MSe $= 66.01$]

The recognition data were analyzed by converting hits and false alarms (for the novel patterns) into d' scores for each subject. Of the 31 subjects receiving the recognition test, all but one recognized the previously seen exemplars at a better than chance rate. Although there was some tendency for d' to decline with delayed testing, this trend was not significant [$F(2,28) = 2.27$, $p > .05$, MSe $= 0.35$]. The recognition and classification data provided converging evidence that subjects retained information specific to the patterns learned during category acquisition (the original category members). For immediate testing, classification accuracy was high for the originals, and recognition accuracy (the ability to discriminate the originals from novel category members) was also relatively high. With the introduction of a two-week delay between acquisition and testing, classification accuracy for the original patterns dropped off substantially, and recognition accuracy was also reduced. A four-week delay brought no further decrements in performance for either classification or recognition. Classification accuracy remained 7% to 9% lower for the novel than the original patterns at all delay intervals between acquisition and testing.

Both the classification and recognition data for the central patterns of each category (determined from the average dot location for each category member presented during acquisition) were very similar to the data obtained for other category members (the novels) that were not presented during acquisition.

Discussion

The results were consistent with those of Experiment 2a in providing evidence for the long-term retention of item-specific information, in this

case over an interval of four weeks. Despite the fact that seven of the
eight dots in each Set I category member were in the same relative
location, the results for the central patterns provided no evidence that
subjects' category representations were based on the averaged locations of
element-level perceptual units.

Experiment 3b

The parsing data obtained in Experiment 1 indicated that the Set III
categories were similar to the Set I categories with regard to the
configural distinctiveness of individual patterns. Also, members of the
same Set III categories, like members of the same Set I categories,
resembled each other at the part-level. However, the similarity data
obtained in Experiment 1 indicated that these resemblances were not as
strong as they were for the Set I categories. Obtaining the typical
original/novel by time-of-test interaction for the Set III categories
would support the hypothesis that both configural distinctiveness and
relatively strong part-level resemblances are required for the long-term
retention of information specific to previously learned dot patterns.

Method

Except for substituting the Set III for the Set I patterns, the design
and procedure were identical to that of Experiment 3a. Thirty-six
undergraduate students at Florida Atlantic University voluntarily
participated in this experiment.

Results

A total of 14.4 trial-blocks were required for category acquisition:
6.2, 5.2, and 3.0 for the first, second, and third training phases,
respectively (again, excluding errorless training blocks).

The classification and recognition data are presented in Table 6.
As for the Set II patterns in Experiment 2b, the original/novel difference
in classification accuracy decreased with increased delay between category
acquisition and classification testing. An analysis of variance indicated
that whether the test patterns were originals or novels interacted with
time-of-test in affecting percent classification accuracy [$F(2,33) = 9.17$,
$p < .001$, MSe = 111.7]. Tests of simple effects indicated that the
original/novel difference was significant when classification testing took
place immediately [$F(1,33) = 30.75$, $p < .001$, MSe = 111.7] or after a
two-week delay [$F(1,33) = 14.00$, $p < .001$, MSe = 111.7], but was not
significant after a four-week delay [$F(1,33) < 1.0$, MSe = 111.7].

- - - - - - - - - - - - - - - -
Insert Table 6 About Here
- - - - - - - - - - - - - - - -

As in Experiment 3a, the recognition data were analyzed by
converting hits and false alarms into d' scores for each subject. All 36
subjects participating in this experiment recognized previously seen
patterns at a better than chance rate. The decrease in their d' scores
with increased delay between acquisition and testing was significant

[F(2,33) = 8.09, p < .001, MSe = .401]. The recognition data were therefore consistent with the trend of the classification data; the ability to discriminate between original and novel category members declined with increasing delay between acquisition and testing. Although information specific to the patterns learned during acquisition was not completely forgotten after four weeks (recognition performance remained better than chance), this information no longer influenced classification performance, which was equivalent for the originals and novels.

The results obtained for the central patterns were difficult to interpret. Classification accuracy for the central patterns was comparable to the original patterns immediately after acquisition and after a two-week delay. After a four-week delay, it was greater than the originals, but the difference was not significant, $t(11) = 1.73$, $p > .05$. For the recognition data, the proportion of "yes" responses for the central patterns was intermediate to the original and novel category members.

Discussion

The classification data obtained for the Set III categories provided the original/novel by time-of-test interaction that is typical of the previous research literature. This result supported the hypothesis that the lack of strong part-level resemblances would impede the long-term retention of information specific to previously learned patterns.

General Discussion

Implicit in most categorization experiments using artificial dot patterns is the assumption that the location of individual elements is the appropriate level of perceptual analysis. As indicated earlier, this assumption has been formalized in conjunction with Knapp and Anderson's (1984) neural summation model. However, the results of this study indicated that perceptual relations among dot patterns belonging to the same category are not adequately represented at the level of individual elements. Despite the fact that seven of the eight dots composing the patterns belonging to the same Set I category were in the exact same relative location, similarity judgments obtained in Experiment 1 provided no evidence that subjects were influenced by the correspondence of element-level perceptual units. Another aspect of our results that was difficult to reconcile with the notion that category relations among dot patterns are represented at the level of individual elements is the original/novel difference in classification accuracy for the Set I patterns. Classification errors might occur for both originals and novels because of difficulty remembering the name for each category, but with seven of eight dots in the same relative location, category representations based on element-level perceptual units would have been sufficiently discriminable for the novels to be classified as accurately as the originals. Our final evidence concerning element-level units was obtained for the central patterns, which were formed by averaging the locations of dots for patterns from the same Set I categories. If subjects' category representations were based on the averaged locations of element-level perceptual units, as specified in models like Knapp and Anderson's (1984), classification performance for the central patterns would have exceeded that obtained for other novel members of the Set I categories. Also consistent with this

conclusion were results reported by Barresi, Robbins, and Shain (1975). Although members of their categories were generated from base patterns that were likely to have common dots, they concluded that the presence of distinctive higher-order features was more important than the presence of common dots in affecting subjects' ability to learn the categories.

Rather than individual elements, the representational structure of categories comprising dot patterns appears to be based on pattern-parts and pattern-configuration, the latter being determined by the spatial arrangement of the parts. Support for this view comes from an analysis of the perceptual units associated with the long-term retention of information specific to previously learned category members. Evidence for long-term retention was obtained only for the Set I categories; the absence of such evidence for the Set II and Set III categories was more typical of the experimental literature based on artificial categories. These results indicated that long-term retention requires both perceptual distinctiveness at the level of pattern-configuration and perceptual resemblance at the level of pattern-parts. Both were present for the Set I categories, but the patterns belonging to the Set II categories lacked configural distinctiveness and the patterns belonging to the Set III categories had relatively weak part-level resemblances. Why are both configural distinctiveness and part-level resemblance necessary for the long-term retention of item-specific information? One possibility is that configural distinctiveness would decrease the likelihood that memory traces of previously learned category exemplars would become indistinguishable from each other as information in the traces is lost over time; part-level resemblance could increase the retrievability of these memory traces by providing organizational "links" among them.

The results of this study therefore indicated that the long-term retention of information specific to previously learned category exemplars that is typical of natural objects can also be obtained for artificial dot patterns, providing their structure reflects the perceptual characteristics that Tversky and Hemenway (1984) identified in their study of natural object categories. The latter argued that configural-level information provides individual category members with a distinctive identity, and part-level information provides the shared attributes that are the basis for determining the membership of objects in basic level categories. The research reported in this paper goes a step further in suggesting that part-level attributes may also contribute to the introduction of organizational "links" that facilitate the retrieval of distinctive, item-specific information.

A final point of interest concerns the consequences of visual similarity for category structure. It is clear from our own data that the relative difficulty of category learning increased as the ratio of within-category to between-category similarity decreased. Mervis and Rosch (1981) have argued previously that the ratio of within-category to between-category similarity is maximized at the basic level of categorization. However, the results of this study suggest that similarity may not be a unitary concept, at least insofar as its affects on categorization are concerned. In the present study, we found that the absence of an original/novel by time-of-test interaction for the Set I patterns was "bracketted" by significant original/novel by time-of-test interactions for

categories whose within-category similarity was greater than (Set II) and less than (Set III) that of Set I (between-category similarity was essentially the same for all three category sets). These results indicated that the ratio of within-category to between-category similarity could not predict whether subjects would retain item-specific information over long periods of time. Long-term retention appears to require both within-category similarity and dissimilarity, but at different levels of perceptual structure.

References

Barresi, J., Robbins, D. & Shain, K. (1975). Role of distinctive features in the abstraction of related concepts. Journal of Experimental Psychology: Human Learning and Memory, 1, 360-368.

Hartley, J. & Homa, D. (1981). Abstraction of stylistic concepts. Journal of Experimental Psychology: Human Learning and Memory, 1, 33-46.

Hock, H. S. & Schmelzkopf, K. F. (1980). The abstraction of schematic representations from photographs of real-world scenes. Memory & Cognition, 8, 543-554.

Homa, D., Sterling, S. & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. Journal of Experimental Psychology: Human Learning and Memory, 7, 418-439.

Knapp, A. G. & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 616-637.

Medin, D. L., Dewey, G. I. & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. Journal of Experimental Psychology: Learning, Memory & Cognition, 9, 607-625.

Mervis, C. B. & Rosch, E. (1981). Categorization of natural objects. Annual Review of Psychology, 32, 89-115.

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. Journal of Verbal Learning and Verbal Behavior, 21, 1-20.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Posner, M. I. & Keele, S. W. (1970). Retention of abstract ideas. Journal of Experimental Psychology, 83, 304-308.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8, 382-439.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. Journal of Experimental Psychology: Human Perception and Performance, 2, 491-502.

Tversky, B. & Hemenway, K. (1984). Objects, parts, and categories. Journal of Experimental Psychology: General, 113, 169-193.

134

Author Notes

Footnotes

[1]Our criteria for determining that the same subunit was shared by (i.e., over-lapped) two patterns was that the subunits had to contain the same number of dots, and they had to be very similar in their relative location within each pattern (within one dot-diameter for the Set I and Set II patterns; within 1.5 dot-diameters for the less compact Set III patterns). Detailed analyses indicated that these relative-location criteria maximized the correlation between similarity judgments and subunit-overlap (part-level resemblance) for pairs of patterns drawn from each set.

[2]The responses of two subjects (in the concept-formation condition) who interchanged the verbal labels for two categories were scored as if the correct labels had been used. This adjustment of the data slightly reduced overall error rates, but had little effect on original/novel differences in classification accuracy.

Table 1

Experiment 1: Mean similarity ratings, within-category and between-category, for Sets I, II, and III. The rating scale extended from 1 (Very Dissimilar) to 10 (Very Similar).

	Set I	Set II	Set III
Within Category	6.7	8.3	5.2
Between Category	3.4	3.7	3.7

Table 2

Assignment of response labels to visual categories for concept-formation and paired-associate training for either the (1-4) or (5-8) subset of category members.

Concept-Formation Training

	Category Members			
Category	1(5)	2(6)	3(7)	4(8)
A	red	red	red	red
B	green	green	green	green
C	blue	blue	blue	blue
D	yellow	yellow	yellow	yellow

Paired-Associate Training

	Category Members			
Category	1(5)	2(6)	3(7)	4(8)
A	red	green	blue	yellow
B	Paris	Rome	London	Madrid
C	penny	nickel	dime	dollar
D	April	June	March	August

Table 3

Experiment 2a: Percent classification errors for the Set I patterns following acquisition with either concept-formation (CF) or paired-associate (PA) training. Classification testing occurred either immediately or after a two-week delay.

Acquisition		Time-of-Test	
		Immediate	Delay
CF	Bases	17.9	17.9
Training	Originals	4.5	10.3
	Novels	13.4	18.8
	Novels minus Originals	8.9	8.5
PA	Bases	12.5	19.6
Training	Originals	2.7	4.5
	Novels	9.4	11.6
	Novels minus Originals	6.7	7.1
mean	Bases	15.2	18.8
	Originals	3.6	7.4
	Novels	11.4	15.2
	Novels minus Originals	7.8	7.8

Table 4

Experiment 2b: Percent classification errors for the Set II patterns following category acquisition with either concept-formation (CF) or paired-associate (PA) training. Classification testing occurred either immediately or after a two-week delay.

| Acquisition		Time-of-Test	
		Immediate	Delay
CF	Bases	1.8	10.7
Training	Originals	1.3	8.0
	Novels	4.5	7.6
	Novels minus Originals	3.2	-.4
PA	Bases	1.8	0
Training	Originals	0.4	1.8
	Novels	2.7	2.7
	Novels minus Originals	2.3	.9
mean	Bases	1.8	5.4
	Originals	.9	4.9
	Novels	3.6	5.2
	Novels minus Originals	2.7	.3

Table 5

Experiment 3a: Percent classification errors, percent "Yes" recognition responses, and d' scores in recognition test for the Set I patterns when classification and recognition testing followed immediately, two weeks, or four weeks after acquisition (concept-formation training).

	Classification Errors			"Yes" Recognition Responses		
	Immediate	2-Weeks	4-Weeks	Immediate	2-Weeks	4-Weeks
Centrals	16.7	24.0	27.1	31.3	66.7	62.1
Originals	3.1	17.2	18.8	75.8	80.7	77.3
Novels	12.0	24.5	27.6	28.9	55.7	66.7
Novels minus Originals	8.9	7.3	8.8	-	-	-
d'	-	-	-	1.37	.87	.88

Table 6

Experiment 3b: Percent classification errors, percent "Yes" recognition responses, and d' scores in recognition test for the Set III patterns when classification and recognition testing followed immediately, two weeks, or four weeks after acquisition (concept-formation training).

|	Classification Errors			"Yes" Recognition Responses		
	Immediate	2-Weeks	4-Weeks	Immediate	2-Weeks	4-Weeks
Centrals	10.4	33.3	33.3	62.5	68.8	60.4
Originals	9.4	28.1	46.9	80.7	81.8	72.9
Novels	33.4	44.3	45.4	23.4	41.1	49.5
Novels minus Originals	9.1	7.3	8.8	-	-	-
d'	-	-	-	1.76	1.25	0.72

Figure Captions

1. Four base patterns and members of the Set I categories.

2. Procedure for generating the Set I category members.

3. Four base patterns and members of the Set II categories.

4. Procedure for generting the Set II category members.

5. Four base patterns and members of the Set III categories.

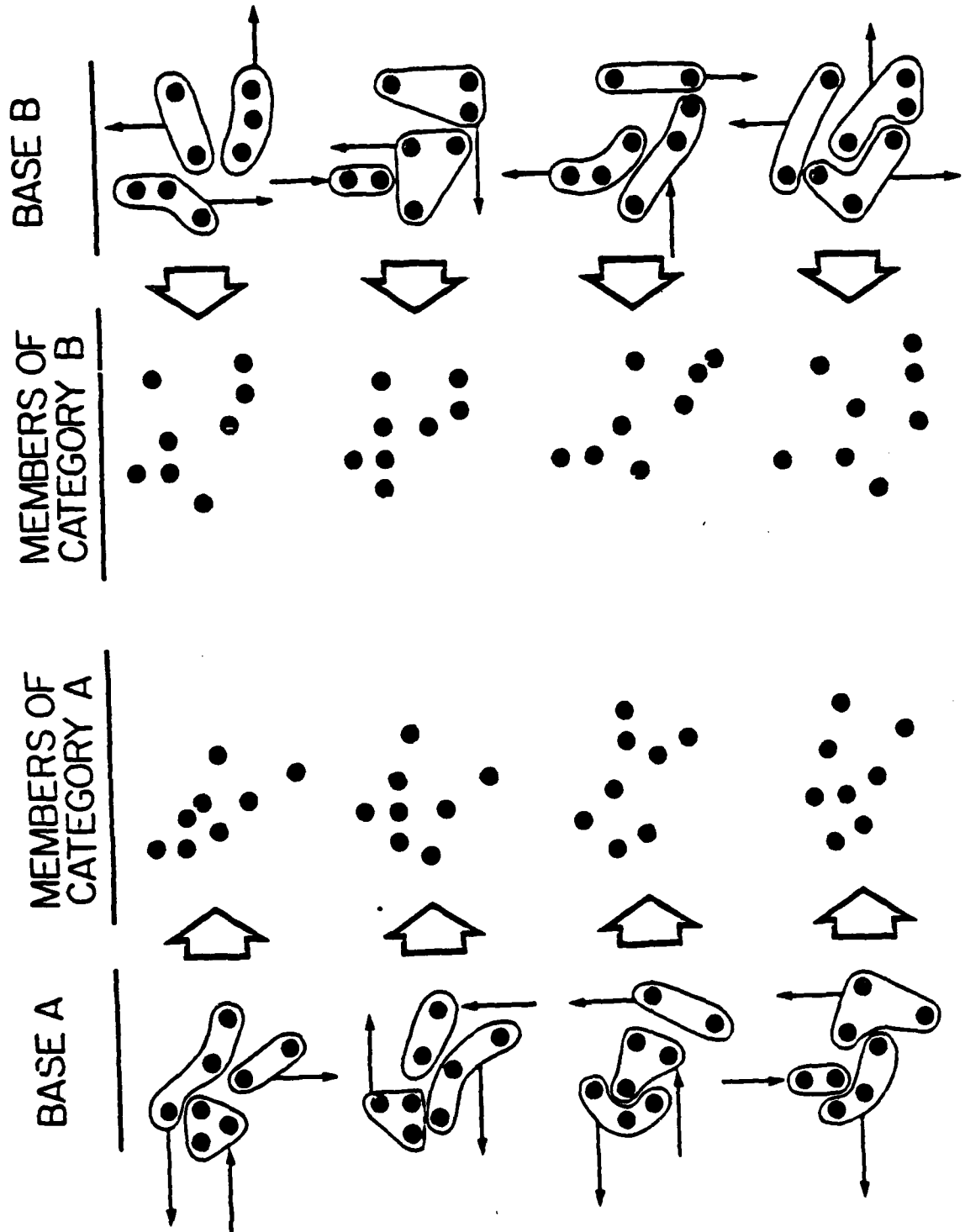6. Procedure for generating the Set III category members.
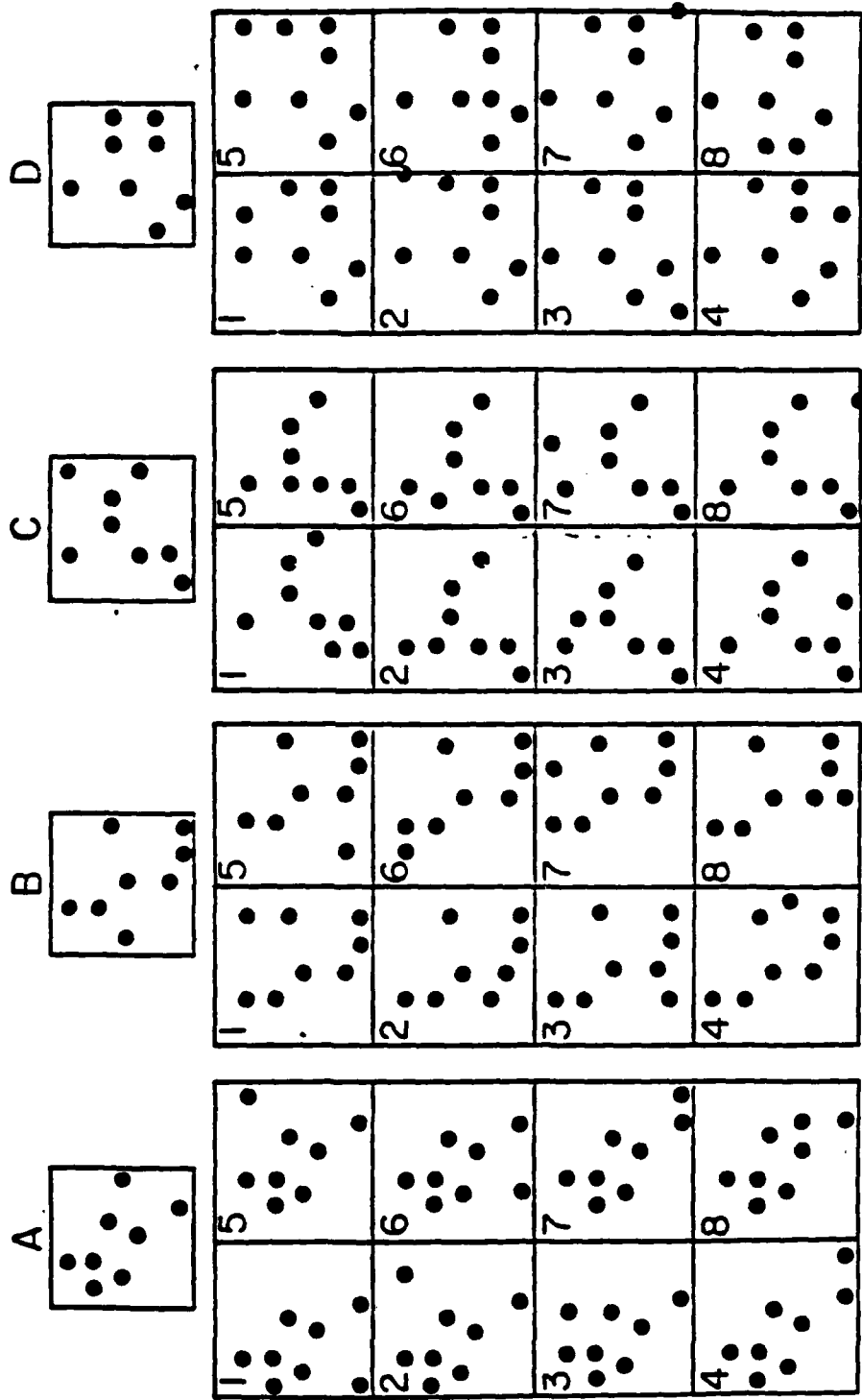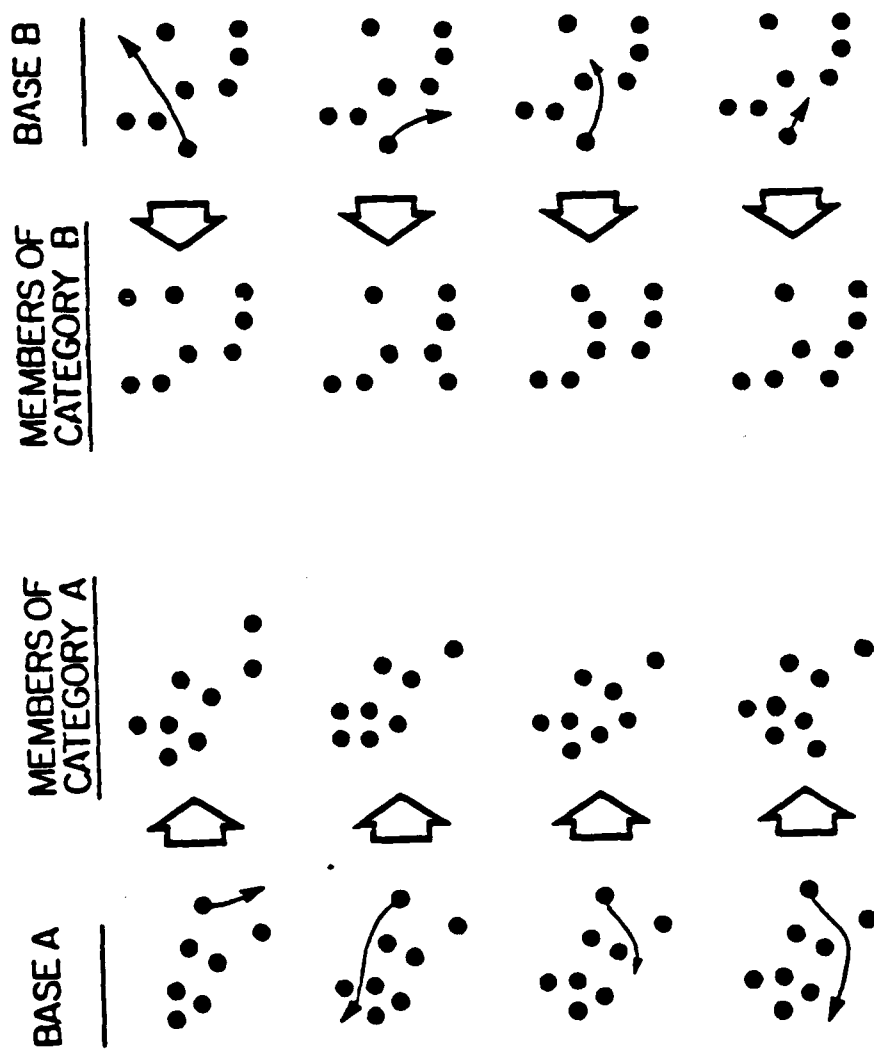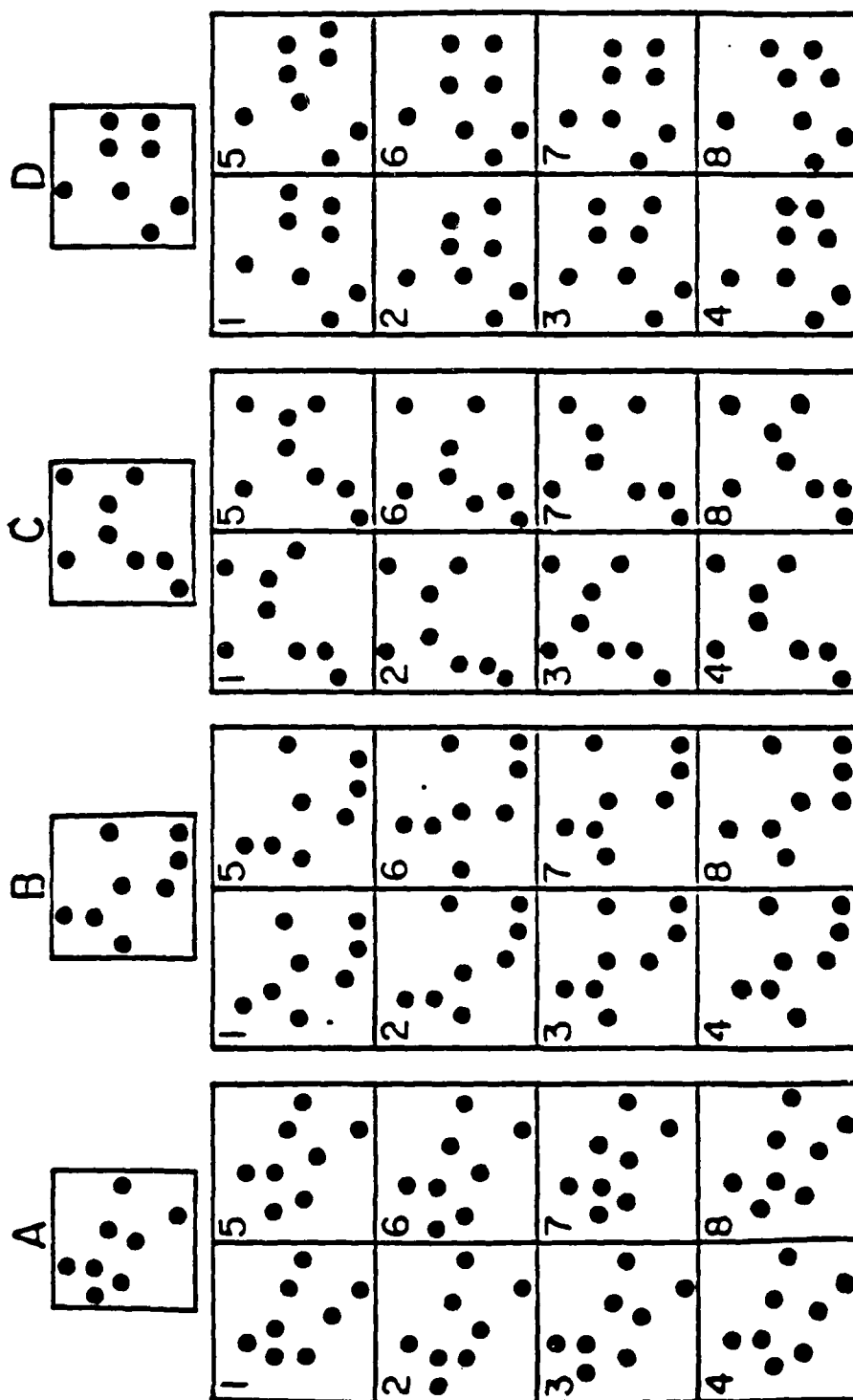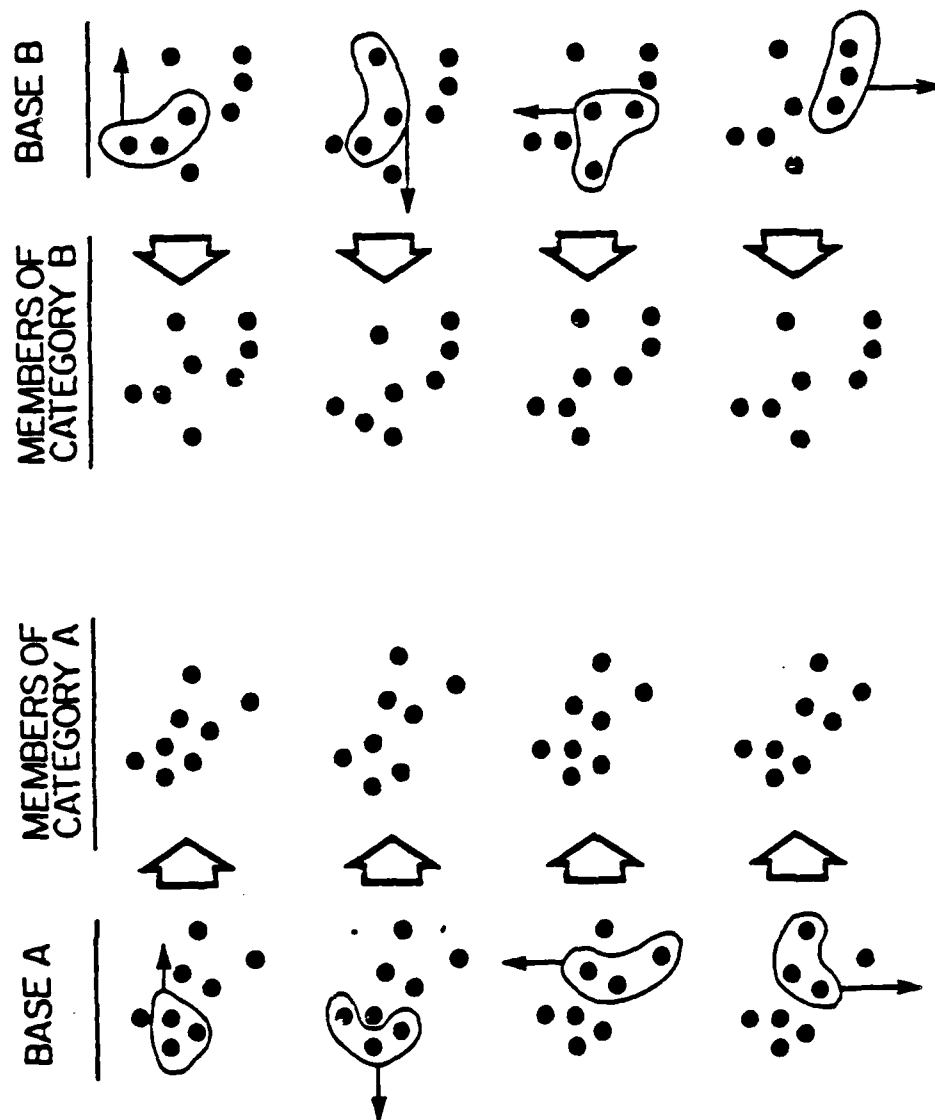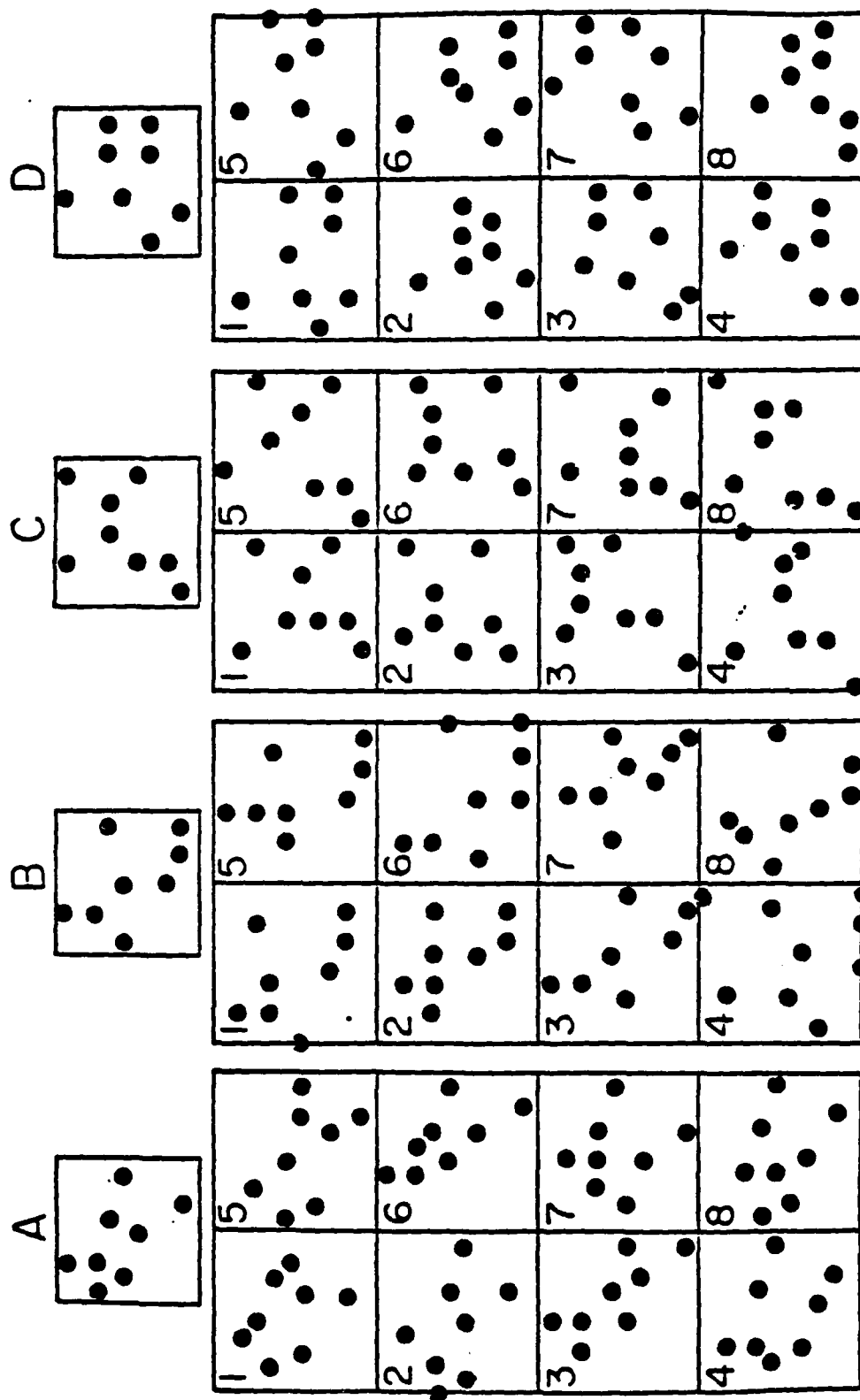
SECTION 3C

Perceptual Learning in Visual Category Acquisition
Howard S. Hock, Elizabeth Webb, and L. Clayton Cavedo

Abstract

The attribute structure of a set of dot patterns was studied by having
subjects segment (parse) the dots of each pattern into parts or subunits.
Subjects identified subunits by drawing circles around groups of dots
from each pattern. These parsing data were obtained for subjects who had
no prior experience with the patterns, and for subjects who had
previously learned to identify the patterns with respect to their
membership in one of four categories. Analyses of the parsing data
indicated that category learning increased the salience of large subunits
that were similar in orientation for patterns that were members of the
same category. This evidence for perceptual learning was obtained even
when the category training procedure required learning to identify the
patterns individually, suggesting that attribute abstraction and
item-learning are not incompatible. It was also obtained despite a
decrease in over-all intersubject agreement. The latter result led us to
question the usefulness of intersubject agreement as an index of category
knowledge.

One of the strategies available to subjects learning the category membership of a set of stimuli is to determine the attributes that are likely to be shared by members of the same category but are unlikely to be shared by members of contrasting categories (i.e., diagnostic attributes). Brooks (1978) refers to this strategy as analytic. The relative difficulty of analytic category learning depends on how readily shared/discriminative attributes can be abstracted (i.e., tested for diagnosticity independent of other attributes).

We would expect most subjects to begin category training analytically by testing attributes that are precategorically salient (a similar assumption is made in Fried and Holyoak's [1984] category density model). These are the attributes that subjects would be most likely to notice and remember in examining the stimuli individually. If the precategorically salient attributes are not diagnostic of category membership, the subject has the option of persisting in an analytic strategy and continuing the search for attributes that are likely to be shared by members of the same category, but not by members of contrasting categories. This would lead the subject to consider, as potentially diagnostic of category membership, attributes that were not the most salient prior to category training. Successful category learning based on searching for shared/discriminative attributes would be accompanied by increases, relative to precategorical values, in the salience of attributes shared by members of the same category (an alternative strategy could result in category learning without changes in attribute salience; Brooks (1978) and Medin, Dewey, and Murphy (1983) have demonstrated that under certain conditions, subjects will forego testing for diagnostic attributes and simply learn to identify stimuli as individuals).

An analytic strategy was also implicit in Eleanor Gibson's (1969) assertion that perceptual learning involves an active search aimed at the discovery of attributes that discriminate among stimuli that are to receive different responses. As for category learning, analytic discrimination learning is likely to begin with a search for attributes that would be highly salient independent of any training. The search would end with the high-salience attributes if they proved to be distinctive to the to-be-discriminated stimuli. Under such circumstances, it would be difficult to empirically demonstrate that subjects were searching for or discovering anything. Evidence that perceptual learning involves searching for and discovering discriminative attributes would be more readily obtained if the analytic strategy persisted, and the search process continued beyond testing the attributes that were most salient prior to training. The subject would then have the opportunity to discover discriminative attributes that had been relatively low in salience prior to training.

Previous research concerned with perceptual learning has focussed on learning to discriminate among stimuli that were initially difficult to tell apart. At issue in this earlier research was whether improvements in discriminability are the result of predifferentiating individual stimuli by associating different responses with each, as argued by Ellis (1973), or whether they are due to increases in the perceiver's sensitivity to distinctive characteristics of the stimuli, as argued by

Gibson and Gibson (1955). The present study differs from these experiments, which were concerned with increases in the salience of discriminative attributes. Instead, it is concerned with perceptual learning that increases the salience of attributes that are shared by members of the same category.

Evidence consistent with the hypothesis that category learning can increase the salience of attributes shared by members of the same category has been obtained by Homa, Rhoads, and Chambliss (1979). They found, as a result of category learning, that pairs of stimuli belonging to the same category increased in judged similarity, whereas pairs of stimuli belonging to contrasting categories decreased in judged similarity. However, Homa, et al.'s (1979) study provided only indirect evidence for perceptual learning; since they were not concerned with perceptual learning, there was no attempt to identify specific attributes that increased in salience as a result of category learning. Also consistent with the hypothesis that category learning can increase the salience of attributes shared by members of the same category are the results of experiments reported by Medin and Smith (1981) and Medin, Dewey, and Murphy (1983). In these experiments, changes in attribute salience were indicated by differences in the parameter values that provided the best fit of Medin and Schaffer's (1978) computational formula to their categorization data. Since different training procedures resulted in different parameter values, it could be inferred that at least one of the training procedures changed the salience of the attributes relative to their precategorical values. However, since Medin and his colleagues were not concerned with perceptual learning, they did not assess the precategorical salience of the attributes. Thus, the relative salience of the attributes they tested may have changed as a result of category learning, but there was no way of determining whether the change constituted an increase in salience without having also assessed the precategorical salience of the attributes.

Although the issue is fundamental to any theory of categorization, there is to our knowledge no published evidence that reports a direct test of whether perceptual learning, in the form of increased salience of attributes shared by members of the same category, can occur as a consequence of category learning. The outcome of such an experiment is by no means obvious. In order to observe an increase in the salience of shared attributes, it is necessary to begin with a set of stimuli for which attention to precategorically salient attributes is not a sufficient basis for successful category acquisition. Having done so, there is no guarantee that there will be an increase, relative to precategorical values, in the salience of attributes shared by members of the same category. As indicated earlier, subjects may test hypotheses regarding the attributes that are precategorically salient, and when that fails, adopt an item-learning strategy instead of searching for nonobvious, diagnostic attributes (Brooks, 1978).

Since previous research has indicated that the nature of the training procedure can influence whether subjects abstract diagnostic attributes during category learning (Brooks, 1978; Medin, et al., 1983), two different acquisition procedures were used in Experiment 1: concept-formation and paired-associate training. The procedures differed in that

only the latter required learning to discriminate among members of the same category. We were interested in whether perceptual learning (i.e., increased salience of attributes shared by members of the same category) would occur with a training procedure that emphasized an item-learning strategy by requiring subjects to learn items as individuals (paired-associate training; each pattern received a different name), as well as a training procedure that emphasized an analytic strategy by instructing subjects to look for relations among the patterns that were relevant to their category membership (concept-formation training; each pattern received the same name). We included classification and recognition tests following category training in order to evaluate performance differences that might be obtained for the two training procedures.

Experiment 1

The stimuli used in this experiment were dot patterns. The attribute structures of these patterns were examined by having subjects segment the patterns into subunits (parts) by drawing circles around groups of dots within each pattern. The parsing procedure we used was similar to procedures used previously by Banks and Prinzmetal (1976) and Palmer (1977). The ability of the latter investigators to predict performance in various tasks on the basis of how their stimuli were segmented into parts indicated that the parsing procedure successfully reflects important aspects of a stimulus' attribute structure. Furthermore, Tversky and Hemenway (1984) have shown that the identifiable parts of an object are critical to the representation of basic level categories for natural objects. These studies demonstrate that the constituent parts of a visual stimulus are empirically and "ecologically" valid units of analysis.

The stimulus attributes that were relevant to our study of perceptual learning were the characteristics of the parts into which the patterns were parsed. The particular characteristics we examined were the number of dots constituting each part, the shape of each part, and the orientation of each part.

The general design of the experiment was to compare how dot patterns were parsed into subunits for three groups of subjects. One group had no category training, one group received concept-formation training, and one group received paired-associate training. It was hypothesized that attributes that were common to patterns which were potential members of the same category (as measured in the no-training condition) would increase in salience when subjects learned to categorize the patterns (as measured in the two training conditions). In order to show that increases in the salience of attributes shared by members of the same category did not occur by chance (if an attribute becomes more salient, its likelihood of re-appearing in the parsing data for many different patterns would also increase), we contrasted the salience of attributes common to patterns drawn from the same category with the salience of attributes common to sets of patterns drawn from different categories. Our method of computing common-attribute salience involved a version of Medin and Schaffer's (1978) multiplicative combination rule. We first measured the salience of an attribute as it appeared in each pattern by

counting the number of subjects who circled the pattern-part when they parsed the pattern. We then determined the common-attribute salience for a set of patterns (i.e., how strongly they resembled each other on the basis of this attribute) by computing the product of the salience values for all the patterns in the set.

Method

Subjects. A total of 192 subjects provided the data for this experiment: 64 were in the no-training condition, 64 received concept-formation training, and 64 received paired-associate training. Some of the 64 subjects in the no-training condition were tested in large groups (they were not paid), others were tested individually (they received $2.00). The 128 subjects in the two training conditions were tested individually and paid $2.00 per hour for their participation in an experimental session lasting between one and one-and-a-half hours. All subjects were undergraduate students at Florida Atlantic University. Their participation was voluntary. They had no previous experience with the patterns or foreknowledge that the research was concerned with category learning.

- - - - - - - - - - - - - - - - - -
Insert Figures 1 and 2 About Here
- - - - - - - - - - - - - - - - - -

Stimuli. Dot patterns were defined as belonging to the same category if they were generated from the same base pattern by the same generation procedure. The four base patterns used to generate the categories are presented in the top row of Figure 1. Each base pattern comprised eight dots whose locations were defined by positions on an imaginary 11 x 11 grid. The spacing between grid lines was equal to one dot-diameter, resulting in patterns that were relatively compact. Category members were generated from the base patterns by partitioning the eight dots composing each base pattern into two groups of three dots and one group of two dots. The dots within each group were then moved one space in the same direction, but the different groups were each moved in different directions. Each pattern generated in this way involved a different partition of the base pattern's eight dots into groups of three, three, and two, and a different pattern of movement-direction for the three groups. The set of eight patterns generated for each base pattern is presented in Figure 1. The procedure for generating the patterns is illustrated in Figure 2. For descriptive convenience, each category will be identified by a letter (A, B, C, or D) and each member of a category by a number (1 through 8). Subset (1-4) of each category comprised Patterns 1, 2, 3, and 4; subset (5-8) of each category comprised Patterns 5, 6, 7, and 8.

Ratings of similarity for all the pairwise combinations (within-category and between-category) of Patterns 1, 2, 7, and 8 of each category were obtained from six subjects who did not participate in any other aspect of the experiment. Their ratings indicated that patterns belonging to the same category resembled each other only slightly more than patterns belonging to different categories. On a scale of one (least similar) to ten (most similar), the mean ratings were 6.6 for

pairs of patterns from the same category and 5.1 for pairs of patterns from different categories.

Design. As indicated earlier, there were three groups of subjects in the experiment. The group assigned to the no-training condition had no experience with the patterns prior to parsing them into subunits; the parsing data constituted their only contribution to the experiment. Since these subjects had no idea that the patterns they saw belonged to categories, their data reflected attributes that were salient independent of the category membership of the patterns. Subjects in the two training conditions participated in an experimental session with three phases: 1) category learning (with either concept-formation or paired-associate training), 2) classification/recognition testing for previously learned and novel category members, and 3) parsing four patterns, each from a different category, into their constituent subunits.

For each of the two training conditions, half the subjects learned to identify the patterns in subset (1-4) of each category, and half learned to identify the patterns in subset (5-8) of each category. Classification/recognition testing was introduced immediately after acquisition. Thirty-six patterns were presented: 16 were the original exemplar patterns that had been presented numerous times during acquisition, 16 were novel category members that had not been seen before, and 4 were the base patterns that also had not been seen before. If subset (1-4) comprised the originals during acquisition, subset (5-8) comprised the novels during classification/recognition testing, and vice versa. The 36 patterns were presented to subjects in one of two random sequences. Matching sequences were provided for the counterbalancing of the (1-4) and (5-8) subsets as originals and novels. Half the subjects received only recognition testing, the other half received a combination of classification and recognition testing (the purpose of this contrast is unrelated to this paper).

Following classification/recognition testing, subjects in the training conditions parsed four patterns that were not members of the acquisition set. The post-training parsing data were obtained for novel category members in order to strengthen the conclusion that changes in attribute salience as a result of category learning would reflect the structure of the categories rather than attributes specific to previously learned patterns. If subjects had been trained with subset (1-4) of each category during the acquisition phase of the experiment, they subsequently parsed Patterns A7, B7, C7, and D7 (in counterbalanced order), or Patterns A8, B8, C8, and D8 (also in counterbalanced order). If they had been trained with subset (5-8) of each category, they subsequently parsed Patterns A1, B1, C1, and D1, or Patterns A2, B2, C2, and D2 (again in counterbalanced order). To insure that the patterns were properly categorized during the parsing phase, the appropriate category label was typed above each of the to-be-parsed patterns (red, green, blue, and yellow for the subjects who had received concept-formation training; city, color, month, and money for the subjects who had received paired-associate training). These labels were not provided for subjects in the no-training condition.

Sixteen subjects provided the parsing data for each of the four packets of patterns (A1B1C1D1, A2B2C2D2, A7B7C7D7, and A8B8C8D8), resulting in a total of 64 subjects in the no-training and 64 subjects in each of the two training conditions (concept-formation and paired-associate training). Subjects in both training conditions were assigned to one of eight subgroups defined by the orthogonal combination of parsing set (A1B1C1D1, A2B2C2D2, A7B7C7D7, or A8B8C8D8) and intervening testing procedure (recognition or classification + recognition). Eight subjects were assigned to each of these eight subgroups.

Procedure. During the acquisition phase of the experiment, black on white slides of the dot patterns were back-projected, via a random-access projector, onto a translucent screen. The size of the patterns on the screen was 2.5 cm x 2.5 cm. They were viewed from a distance of approximately 114 cm, thereby intercepting a visual angle of 1.3 deg. Subjects receiving concept-formation training learned the same verbal label for each member of the category (see Table 1). They were told that they would be seeing 16 dot patterns, each of which belonged to one of four groups, and were instructed to look for relations among the patterns that would allow them to learn which patterns belonged in each group. In paired-associate training, subjects learned different verbal labels for each pattern (see Table 1). However, the names for members of the same visual category belonged to the same semantic category. The labels for Category A were all names of colors, Category B names of cities, Category C names for money, and Category D names of months. Subjects receiving paired-associate training did not receive any instruction directing them to look for relations among the patterns (although the semantic similarity of the response labels for members of the same category could have had the same effect).

```
- - - - - - - - - - - - - - - - - -
          Insert Table 1 About Here
- - - - - - - - - - - - - - - - - -
```

Acquisition began with subjects being shown 16 exemplar patterns in random order. One group was trained with the (1-4) subset of each category, the other with the (5-8) subset of each category. The experimenter provided the label for each pattern orally during this initial presentation. Subsequent blocks of 16 trials were presented in one of five random sequences. The order of these trial-block-sequences was varied from subject to subject. Subjects were given up to 10 seconds to verbally identify each pattern, at which point they were required to make their best guess. Corrective feedback was provided following incorrect responses. The acquisition phase proceeded in this manner until two consecutive blocks of 16 trials were completed without error.

Classification/recognition testing followed immediately after category training. Subjects were told that they would be seeing some old patterns and some new ones they hadn't seen before, and that the patterns belonged to the same groups that they had already learned. Those receiving only recognition testing were required to indicate whether or not they had previously seen each pattern. Those receiving a combination. of classification and recognition testing had to first categorize each

test pattern and then indicate whether or not they had seen it before (they were provided with the appropriate category names on an index card) Each pattern was presented, again with a random-access slide projector, for as long a period of time as subjects needed to respond. They were required to guess if they were unsure of the correct response(s). No corrective feedback was provided.

During the parsing phase, each subject was given a single booklet comprising four 14 cm x 20 cm sheets of white paper. The four patterns parsed by each subject came from different categories in order to avoid the possibility that subjects would feel constrained to repeat parsings should they recognize a pattern as similar to a previously parsed pattern. One pattern (2.5 cm x 2.5 cm) was presented on each sheet.

Parsing instructions. Subjects were instructed to draw circles around the dots in each pattern to form three clusters (the number of clusters was restricted so that each subject's data would have equal weight in determining the perceptual subunits for each pattern). Subjects who parsed the patterns after either concept-formation or paired-associate training were told that their clusters should reflect the ways in which each pattern resembled other members of its category. Subjects in the no-training condition were told that their clusters should reflect their perception of the natural groupings of dots within the pattern. For both the no-training and training conditions, the clusters subjects formed could overlap, and one cluster could be contained within another. Every dot had to appear in at least one circle, even if it appeared alone, and no circle could surround all the dots in the pattern.

Parsing analysis. The perceptual subunits (i.e., the pattern-parts) circled by subjects in parsing each pattern were analyzed in the same way for subjects in the no-training and training conditions. Each subject-selected subunit was coded according to salience, number, orientation, shape, and location. These attributes were not considered exhaustive of all reasonable possibilities.

The salience for each subunit (part) was determined by the frequency with which the subunit was circled. Number refers to the number of dots within the subunit (from one through seven). Based on our assumption that it would be a psychologically important attribute only when the subunit was relatively large, the orientation of the subunit was coded only when the two most distant dots were separated by a distance of at least six dot-diameters (which was more than half of every pattern's height and width). The orientation, expressed in degrees, was computed from the slope of the line defined by the two most distant dots in the subunit. The alternative shape codes used to describe each subunit included the following: linear (vertical, horizontal, left-diagonal, and right-diagonal), triangular (obtuse, acute, and right), parallelogram, rectangular, and "Y" shaped arrangements. The location code for each subunit was determined as follows: 1) the centroid was computed for the pattern comprising the subunit; this was the average x,y location of the eight dots composing the pattern, and 2) the relative location of each subunit within its pattern was computed as the average x,y distance of the dots composing the subunit from the centroid of the entire pattern.[1]

Our analysis of common-attribute salience involved selecting different sets of four patterns (some from the same category, some from different categories), and determining whether there was an attribute-match among the subunits of each of the four patterns. This was done independently for the attributes of number, orientation, and shape; when subunits of different patterns were compared for one of these attributes, the values of the other attributes were ignored. Matches based on number were obtained when there were four subunits, one in each of the four patterns, that had the same number of dots. Matches based on orientation were obtained when there were four subunits, one in each of the four patterns, that had orientations within 45 deg of each other. Matches based on shape were obtained when there were four subunits, one in each of the four patterns, that had the same shape code.

The location of subunits within each pattern introduced the problem of criterion. For example, if each pattern in a set of four patterns had a subunit with the same shape, how close in location would the subunits have to be in order to be considered as matching? Our solution to this problem was to systematically vary the criterion used to determine whether matching subunits in a set of patterns were in the same relative location within their respective patterns. Thus, the presence of matching subunits was tested when the subunits had to differ in relative location by less than one dot-diameter, by less than two dot-diameters, and so on, in order to be considered as matching in either number, orientation, or shape. When subunit matches were obtained for one of these attributes, and the four matching subunits were in the same relative location within their patterns (based on the relative-location criterion), common-attribute salience was computed by multiplying the frequency with which each of the four matching subunits was circled when subjects parsed the four patterns. For example, if four of the patterns belonging to Category A (Patterns 1, 2, 7, and 8) had a rectangular subunit at the same relative location, the salience of this common attribute was determined by computing the product of how frequently the subunit was circled for each pattern ($f_1 \times f_2 \times f_7 \times f_8$). Since the same four patterns could have another shape in common (e.g., subunits forming right triangles), the total common-attribute salience for the attribute of shape was the sum of the frequency-products for each shape that was common to all four patterns. Finally, the fourth root of the summed frequency-products was computed in order to place the computation of common-attribute salience on a scale that would reflect the relative salience of the matching subunits within each of their patterns (the maximum value for each subunit, 16, would be obtained if the subunit common to all four patterns was circled by all 16 subjects for each of the patterns).

The above analysis was performed for eight sets of four patterns. The analysis for the first four sets, A1A2A7A8, B1B2B7B8, C1C2C7C8, and D1D2D7D8, assessed common-attribute salience for patterns belonging to the same category (recall that each subject parsed only one pattern from each of the four categories). The analysis for the second four sets, A1B2C7D8, A2B7C8D1, A7B8C1D2, and A8B1C2D7, assessed common-attribute salience for patterns belonging to different categories. The latter four sets of patterns represent only a sample of all possible between-category combinations. In selecting this sample, we avoided between-category combinations corresponding to the packets of four patterns presented to

subjects in the parsing task (A1B1C1D1, A2B2C2D2, A7B7C7D7, and A8B8C8D8). For each attribute tested, we contrasted within-category and between-category common-attribute salience for the no-training, concept-formation training, and paired-associate training conditions.[2]

Results

A number of different data analyses are reported in this section. The primary analysis is of the parsing data, which provides our basic evidence for the effect of category learning on the salience of attributes shared by patterns belonging to the same category. Preceding the analysis of the parsing data are analyses of acquisition rates, post-acquisition recognition accuracy, and post-acquisition classification accuracy.

- - - - - - - - - - - - - - -
Insert Figure 3 About Here
- - - - - - - - - - - - - - -

Acquisition. All 128 subjects in the training conditions reached our criterion of two errorless blocks of 16 trials. Subjects receiving concept-formation and paired-associate training required an average of 14.4 and 8.0 trial-blocks (excluding the two errorless criterion blocks), respectively, to reach the acquisition criterion. The advantage of paired-associate training was statistically significant [t(126) = 5.46, p < .001]. We also compared the two training conditions with regard to the proportion of errors within each block of 16 trials. In order to control for differences in response confusability in the two training conditions (there were 16 different responses in the paired-associate condition, compared with 4 in the concept-formation condition), the paired-associate data were scored only for between-category acquisition errors (e.g., if a subject responded "nickel" to a pattern that should have been called "dime," they were told that their response was wrong, but we subsequently scored it as correct). The mean error percentage rates per trial block (all between-category errors) for the two training conditions are presented in Figure 3. An analysis of variance based on the first 10 trial-blocks indicated that fewer errors were made in paired-associate compared with concept-formation training [F(1,126) = 506.54, p < .001, MSe = 1568.87]. The significant interaction between training procedure and trial-block [F(9,1134) = 7.04, p < .001, MSe = 114.84] reflected the emergence of the paired-associate advantage in category learning after the first block of acquisition trials. Because the two training procedures differed with regard to the need for individual item learning, it could be concluded that the paired-associate advantage resulted from enhanced item learning. The proportion of all errors during paired-associate training that were within the same category (e.g., calling a pattern "Rome" when the correct response was "Madrid") started at 29% for the first block of trials, and gradually increased as practice proceeded. This rate of within-category errors was consistently higher than the 20% rate that would have occurred strictly by chance.

- - - - - - - - - - - - -
Insert Table 2 About Here
- - - - - - - - - - - - -

Classification/recognition testing. The results of this testing phase, which was interposed between the training and parsing phases of the experiment, are summarized in Table 2. The recognition data were analyzed by converting hits and false alarms into d' scores for each subject. All 128 subjects receiving the recognition test discriminated novel category members from previously seen category exemplars at a better than chance rate. A comparison of the two training procedures indicated that d' scores (based on false alarms for the novel patterns) were significantly better for subjects receiving paired-associate training (mean = 2.9) compared with subjects receiving concept-formation training (mean = 2.5), [$t(126) = 3.35$, $p < .002$]. Subjects receiving paired-associate training also made significantly fewer false recognition responses on the base patterns than subjects receiving concept-formation training [$t(126) = 3.54$, $p < .001$]. It should be noted that superior recognition performance was obtained for the paired-associate condition even though subjects in the paired-associate condition received fewer exposures to the original (exemplar) patterns during training than subjects in the concept-formation condition.

The classification data were very similar for the concept-formation and paired-associate conditions. An analysis of variance on percent classification errors indicated that the original/novel difference was significant [$F(1,62) = 202.95$, $p < .001$, MSe = 130.85], but the effect of training procedure [$F(1,62) < 1.0$, MSe = 187.49] and the interaction between pattern type (original/novel) and training procedure [$F(1,62) < 1.0$, MSe = 130.85] were not significant. Although classification accuracy was much lower for the novel than the original category members, subjects' classification of the novels (33.2% error rate) was nonetheless much better than chance (75% error rate). All 64 subjects classified the novels at better than chance rate. The significance of these data is that when subjects went on to the parsing phase of the experiment, their ability to categorize the novel patterns was very similar following concept-formation and paired-associate training.

Parsing. Our initial examination of the parsing data was concerned with the extent to which there was agreement regarding the subunits circled by each subject. We assessed intersubject agreement by counting the number of different subunits circled by all 16 subjects parsing each pattern and comparing this to the number of different subunits that would have been circled had there been no agreement among the subjects (3 x 16 = 48). If all subjects were in complete agreement, the same three subunits would have been circled by each subject, giving an agreement factor of 3/48 = .06; no agreement would be indicated by an agreement factor of 48/48 = 1.00. We found that category learning resulted in decreased levels of intersubject agreement relative to the no-training condition. The agreement factor was .35 for subjects in the no-training condition, .45 for subjects receiving concept-formation training, and .46 for subjects receiving paired-associate training. The increase in intersubject variability was consistently observed for all 16 patterns that were tested; a reversal was obtained for only one pattern, and then only following concept-formation training.

As indicated in the Method section, our criterion for determining whether a subunit was large enough for its orientation to be a

psychologically important variable was that the two most distant dots in the subunit were separated by at least six dot-diameters (more than half the height and width of every pattern). In comparing the parsing data for the no-training and training conditions, we found that large subunits were more salient for subjects in the training conditions compared with subjects in the no-training conditions. The percentage of subunits that were large by this criterion were 19.0%, 30.9%, and 31.0%, for the no-training, concept-formation, and paired-associate conditions, respectively. With individual patterns serving as the random variable, an analysis of variance indicated that the effect of training procedure (no-training, concept-formation training, paired-associate training) on the percentage of large subunits in a pattern was significant [$F(2,24)$ = 15.70, $p < .001$, MSe = 48.22], and further, that the interaction between the training procedure and the category to which the patterns belonged (A,B,C,D) was not significant [$F(6,24) < 1.0$, MSe = 48.22]. That is, the effect of training procedure on subunit size was similar for all four categories used in the experiment. Newman-Keuls comparisons indicated that the physical size of subjects' subunits was significantly larger following either concept-formation or paired-associate training compared to the no-training condition, $p < .05$.[3]

- - - - - - - - - - - - - -
Insert Figure 4 About Here
- - - - - - - - - - - - - -

In addition to increasing the salience of large subunits, category learning increased the likelihood that patterns belonging to the same category would share large subunits that were similar in orientation. The latter was assessed in terms of "common-attribute salience." The computation of orientational common-attribute salience for a set of patterns required that all the patterns in the set contain large subunits with orientations within 45 deg of each other. The computations were performed for sets of four patterns (e.g., A1A2A7A8) from the same category, and sets of four patterns (e.g., A1B2C7D8) from different categories (see Method section for details). As can be seen in Figure 4, common-attribute salience increased in all conditions as the relative-location criterion was relaxed. The looser relative-location criteria allowed more subunits to be considered as matching than the more stringent criteria. Figure 4 also indicates that common-attribute salience increased, relative to the no-training condition, as a result of both concept-formation and paired-associate training. Increases were observed for sets of four patterns from the same category (within-category) as well as sets of four patterns from different categories (between-category). The increase in between-category common-attribute salience as a result of category learning could be attributed to the increased prevalence of large subunits, which increased the likelihood of large subunits of similar orientation occurring by chance. Most importantly, common-attribute salience for patterns from the same category was greater than the baseline/chance level observed for patterns from different categories for subjects in the two training conditions, but there was no difference from baseline for subjects in the no-training condition.[4]

An analysis of variance was performed in which the random variable was the set of four patterns analyzed for orientational common-attribute

salience. There were four sets of four patterns in the within-category condition (A1A2A7A8, B1B2B7B8, etc.) and four sets of four patterns in the between-category condition (A1B2C7D8, A2B7C8D1, etc.). The analysis of common-attribute salience indicated that the interaction between training procedure (no-training, concept-formation training, paired-associate training) and the type of pattern set (within-category vs. between-category) was significant [$F(2,18) = 16.38$, $p < .001$, MSe = 5.18]. This interaction was not significantly affected by the relative-location criterion used to determine whether there were matching subunits within each set of four patterns tested. That is, the three-way interaction between training procedure, type of pattern set, and relative-location criterion was not significant [$F(16,144) < 1.0$, MSe = 0.23]. Tests of simple effects indicated that the effect of type of pattern set (within-category vs. between-category) on common-attribute salience was significant following concept-formation training [$F(1,18) = 15.11$, $p < .002$, MSe = 5.18], and was also significant following paired-associate training [$F(1,18) = 10.13$, $p < .02$, MSe = 5.18], but was not significant in the no-training condition [$F(1,18) < 1.0$, MSe = 5.18]. Finally, the overall effect of training procedure on common-attribute salience was significant [$F(2,18) = 22.42$, $p < .001$, MSe = 5.18], with subsequent Newman-Keuls comparisons indicating that common-attribute salience was not statistically different in the concept-formation and paired-associate conditions, $p > .05$, and also, that both training conditions resulted in higher levels of common-attribute salience than the no-training condition, $p < .05$.

Discussion

Although large subunits increased in salience as a result of category learning (relative to the no-training condition), we did not consider this sufficient evidence for perceptual learning. The increased salience of large subunits could have been due to a general learning strategy elicited by our category training procedures; there may have been a bias to emphasize large pattern-parts. Stronger evidence for perceptual learning would be obtained if it could be demonstrated that increases in attribute salience were specific to the pattern information acquired during category training. Our primary evidence for perceptual learning was therefore based on the orientational similarity of subunits (parts) of patterns belonging to the same category. As a result of category training, we obtained an increase in the salience of large subunits that were similar in orientation for patterns belonging to the same category (a comparable increase was not obtained for patterns belonging to different categories). In concluding that we have evidence for perceptual learning, we are not arguing that subjects' perception of the natural grouping of elements in a pattern has changed. Rather, our claim is that we have detected an attribute, common to members of the same category, whose likelihood of being noticed increased as a result of category learning.

Experiment 2

A possible limitation in interpreting the results of Experiment 1 concerned the no-training condition. In contrast to subjects in the concept-formation and paired-associate conditions, subjects in the no-training condition parsed the patterns without knowing that they were

potentially members of contrasting categories; they were told that the
pattern-parts they circled should reflect their perception of natural
groupings of dots within each pattern. Our objective in choosing the
instructions for the no-training condition was to assess attribute
salience when each pattern was treated as an individual, independent of
its relationship with any other pattern. It could be argued, however,
that the way subjects parsed the patterns in this condition was peculiar
to our instruction concerning "natural groupings of dots," and had nothing
to do with their lack of knowledge (or training) concerning the category
membership of the patterns. That is, the no-training parsing data might
not have differed from the post-training parsing data if we had instructed
our no-training subjects differently.

In this experiment, subjects again parsed the patterns without
knowing that they were members of different categories (recall that each
subject parsed only four patterns, one from each of the four categories).
However, in contrast with the no-training condition of Experiment 1,
subjects were instructed to circle groups of dots in each pattern that
made the four patterns in each packet look different from each other.[5]
In this way, the instructions focussed on a potentially important aspect
of the relationship among our patterns, the attributes that could
potentially discriminate among member of contrasting categories, but
without introducing the extensive category training provided by the two
post-training conditions of Experiment 1. We measured the level of
intersubject agreement, determined the salience of large subunits, and
assessed orientational common-attribute salience for sets of four patterns
from the same category and sets of four patterns from different
categories.

Method

The patterns used in this experiment were identical to those used in
the parsing phase of Experiment 1. The packets of four patterns given to
subjects and the general instructions regarding parsing the patterns into
clusters were the same as in Experiment 1. The only difference in
procedure compared with the no-training condition of Experiment 1
concerned the instructions. As indicated above, subjects were instructed
to circle groups of dots in each pattern that made the four patterns in
each packet look different from each other. They were told to examine all
the patterns before beginning to circle groups of dots for the first
pattern in the packet. Sixty-four subjects, all tested in large groups,
provided the data for this experiment. All were undergraduate students at
Florida Atlantic University who had no previous experience with the
patterns or foreknowledge that the research was concerned with category
learning.

Insert Figure 5 about here

Results

Intersubject agreement was similar to that observed for the training
conditions of Experiment 1 (.44 in this experiment compared with .45 and

.46 for the concept-formation and paired-associate conditions of Experiment 1). The salience of large subunits was similar to that obtained in the no-training condition of Experiment 1; 21.2% of the subunits were relatively large, compared with 19.0% in the no-training condition of Experiment 1. Over-all levels of orientational common-attribute salience were greater than the levels observed for the no-training condition of Experiment 1, but there was no difference in common-attribute salience for sets of patterns from the same category and sets of patterns from different categories (see Figure 5). An analysis of variance indicated the type of pattern set (within-category vs. between-category) did not significantly affect common-attribute salience [$F(1,6) < 1.0$, MSe $= 4.94$]. Furthermore, this result was not influenced by the relative-location criterion; the interaction between the type of pattern set and the value of the relative-location criterion was not significant [$F(8,48) < 1.0$, MSe $= 0.22$].

Discussion

Although large subunits were no more salient in this experiment than in the no-training condition of Experiment 1, over-all levels of orientational common-attribute salience were greater in this experiment (i.e., the large subunits were more likely to be similar in orientation). However, the absence of a difference in common-attribute salience for sets of patterns from the same category and sets of patterns from different categories indicated that the relatively high level of common-attribute salience (compared with the no-training condition of Experiment 1) was not meaningfully related to the category structure of the patterns. We could conclude, therefore, that the results obtained in the no-training condition of Experiment 1 did not depend on our use of instructions that treated each pattern as an individual, independent of its relationship with any other pattern. Even with instructions emphasizing the relationship among the four patterns being parsed (the differences among them), there was no indication that the orientation of large subunits was a precategorically salient attribute that was more likely to be shared by members of the same category compared with members of different categories.

Experiment 2 was similar to the no-training condition of Experiment 1 in that every subject parsed four patterns, one from four different categories. The difference from the first experiment was in the instructions given to subjects prior to parsing the patterns. In Experiment 3, which follows, subjects again parsed four patterns, but now each pattern belonged to the same category. This allowed us to use instructions similar to the parsing instructions that were used following category training in Experiment 1. That is, subjects were told that their clusters should reflect ways in which the four patterns were similar to each other (subjects in the training conditions of Experiment 1 had actually learned the categories and were therefore instructed to form their clusters to reflect ways in which each pattern resembled other members of its category). The instructions for Experiment 3 therefore focussed on another potentially important aspect of the relationship among the patterns, the attributes that were potentially shared by members of the same category, again without introducing the extensive category training provided by the two post-training conditions of Experiment 1.

Experiment 3

Method

The patterns used in this experiment were identical to those used in the previous experiments. However in this experiment, the packets of four patterns given to subjects all came from the same category. There were 16 packets comprising patterns A1A2A7A8, 16 comprising patterns B1B2B7B8, 16 comprising patterns C1C2C7C8, and 16 comprising patterns D1D2D7D8. The four patterns in each set were presented in four different, counter-balanced orders. As indicated above, subjects were instructed to circle groups of dots in each pattern that made the four patterns in each packet look similar to each other. They were told to examine all the patterns before beginning to circle groups of dots for the first pattern in the packet. Sixty-four subjects, all tested in large groups, provided the data for this experiment. All were undergraduate students at Florida Atlantic University who had no previous experience with the patterns or foreknowledge that the research was concerned with category learning.

Insert Figure 6 about here

Results

The parsing data indicated that the level of agreement regarding the subunits circled by each subject (the agreement factor was .41) was intermediate to the agreement factors obtained in the no-training and training conditions of Experiment 1. The salience of large subunits was again similar to that obtained in the no-training condition of Experiment 1; 21.7% of the subunits were relatively large, compared with 19.0% in the no-training condition of Experiment 1. Despite the relatively low salience of large subunits, over-all levels of orientational common-attribute salience were similar to those obtained for the training conditions of Experiment 1, and common-attribute salience was greater for sets of patterns from the same category compared with sets of patterns from different categories (see Figure 6). This difference, however, depended entirely on the common-attribute salience obtained for patterns from one of the four categories (Category C). Excluding the latter, there was no difference in common-attribute salience for patterns from the same and different categories. The absence of reliability across items was reflected in the analysis of variance, which indicated that the type of pattern set for which attribute-matches were obtained (within-category vs. between-category) did not significantly affect common-attribute salience [$F(1,6) = 1.34$, $p > .05$, $MSe = 20.10$]. This result was not significantly influenced by the relative-location criterion; the interaction between the type of pattern set and the value of the relative-location criterion was not significant [$F(8,48) < 1.0$, $MSe = 0.55$].

Discussion

Large subunits were no more salient in this experiment than in Experiment 2 or the no-training condition of Experiment 1. Over-all levels of orientational common-attribute salience were again relatively

large compared with the no-training condition of Experiment 1, but orientational common-attribute salience for patterns from the same category was not reliably greater than the baseline/chance level observed for patterns from different categories. It was not surprising that the analysis of common-attribute salience yielded results that resembled (though unreliably) those obtained for the training conditions of Experiment 1. By providing each subject with four patterns from the same category, and asking them to look for similarities among them, we were in effect providing them with an opportunity to learn something about the categories. Nonetheless, large subunits remained relatively low in salience and similarities in subunit orientation appeared to be noticed only for patterns from one category.

We have therefore supplemented the no-training condition of Experiment 1 with two conditions in which subjects were instructed to emphasize relationships among the four patterns that they parsed; one condition emphasized differences among potential members of contrasting categories (Experiment 2), the other similarities among potential members of the same category (Experiment 3). Both of these relationships were relevant to analytic category learning, which requires the abstraction of attributes that are more likely to be shared by members of the same category (similarities) than members of contrasting categories (differences). The results of these experiments therefore indicated that large subunits (pattern-parts) were not precategorically salient, and there was no precategorical disposition for the orientation of large subunits to be more similar for sets of patterns that were to become members of the same category compared with sets of patterns that were to become members of different categories. We can conclude, therefore, that the category learning provided by the concept-formation and paired-associate training procedures of Experiment 1 was responsible for the increased salience of large subunits that were similar in orientation for patterns belonging to the same category.

General Discussion

Our perceptual learning hypothesis was framed in conjunction with an analytic category learning strategy, which emphasizes the search for attributes that are diagnostic of category membership (Brooks, 1978). Evidence for perceptual learning was obtained under training conditions that emphasized analytic category learning (the concept-formation condition), but was also obtained under training conditions which appear to have induced an item-learning strategy (the paired-associate condition). Evidence for the latter came from an examination of the acquisition and classification/recognition data for the two training conditions of Experiment 1.

The acquisition data indicated that paired-associate training resulted in faster category learning than concept-formation training, despite the former having a more complex stimulus-response mapping than the latter. This result was consistent with previous research indicating that when stimuli belonging to the same category are sufficiently dissimilar, acquisition can benefit from a training procedure that requires learning the stimuli as individual items (Brooks, 1978; Medin et al., 1983). Better recognition accuracy (original/novel discrimination)

following paired-associate compared with concept-formation training provided further evidence that the former was more facilitative of item learning. Despite these differences, subjects in the two training conditions were alike with regard to their ability to classify previously seen and novel category members. They were also alike with regard to their parsing of the patterns. For both training procedures, patterns belonging to the same category increased in their tendency to share large subunits that were similar in orientation (relative to the no-training condition). Evidence for the abstraction of shared attributes was therefore obtained in both training conditions, even though subjects appear to have adopted an item-learning strategy in the paired-associate condition.

Although Brooks (1978) characterized abstractive strategies as analytic and item-learning strategies as nonanalytic (and presumably nonabstractive), the results obtained in the paired-associate condition indicate that the abstraction of attributes shared by members of the same category is not empirically incompatible with item-learning. In another study, Hock, Tromley, and Polmann (1986) argued that the long-term retention of previously learned category exemplars may be functionally dependent on the abstraction of shared attributes; attributes shared by members of the same category could facilitate the retention of individual patterns by providing organizational "links" among the memory representations corresponding to the previously seen category members. Also, Hock, Smith, Escoffery, and Bates (1986) found that superficial pattern details that might be expected to be encoded in a nonanalytic, pictorial format, are abstracted from patterns in the same way as attributes shared by patterns belonging to the same category.

The results of these experiments, together with those of the present study, suggest that category learning might always be analytic. The category learner may abstract attributes shared by members of the same category and/or attributes that are unique to particular category members. Differences in within-category and between-category similarity, as well as differences in training procedure, could influence the extent to which the category learner abstracts shared vs. item-specific attributes.[6] Furthermore, both the shared and distinctive attributes could be stored in exemplar format (as argued by Medin and Schaffer, 1978), or the shared attributes could be stored in separate, central representations (as maintained by investigators going back to Bartlett, 1932). In either case, the information retained would be in the form of abstracted attributes.

Our concluding discussion concerns the issue of intersubject agreement. The results of the present study indicate that perceptual learning took place in the absence of increases in intersubject agreement. This result is of interest because levels of intersubject agreement on the attributes that characterize various objects or concepts has emerged in the literature as an important empirical index of category knowledge (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Bream, 1976; Tversky & Hemenway, 1983, 1984; Murphy & Wright, 1984; Rifkin, 1985). Murphy and Wright's (1984) paper is particularly relevant to the present study. They report the results of an attribute-listing experiment for three diagnostic categories of childhood adjustment problems; subjects

were required to provide verbal lists of the attributes of each adjustment problem. Instead of studying the acquisition of these categories, they compared the attribute listings generated by individuals with varying levels of real-world expertise (ranging from supervisors in a residential treatment program to students in undergraduate psychology courses). Murphy and Wright (1984) found that intersubject agreement on the characterisic attributes of these disorders increased with higher levels of expertise, but the extent to which these attributes differentiated among the three categories actually <u>decreased</u> with expertise.

Murphy and Wright's (1984) results suggest that there may be a problem with the assumption, implicit in experiments using the attribute-listing technique, that intersubject agreement is the hallmark of category knowledge. Central to this assumption is the expectation that category learning should result in increased intersubject agreement concerning the attributes that are diagnostic of category membership. However, Murphy and Wright (1984) found that increased intersubject agreement was not accompanied by increased salience of attributes that discriminate among contrasting categories, and we have found in the present study that an increase in the salience of attributes shared by members of the same category was obtained without an increase in intersubject agreement.

A critical factor affecting changes in intersubject agreement as a result of training may be the number of attributes that are potentially diagnostic of category membership. If intersubject agreement is to increase as a result of category learning, then the number of potentially diagnostic attributes for the category must be relatively small. The larger the number of potentially diagnostic attributes for category learners to choose among, the greater the likelihood that they will disagree regarding the particular attributes they select as diagnostic of category membership. For the dot patterns used in the present study, the number of attributes that were potentially diagnostic of category membership was likely to have been quite large. Each dot, each pair of dots, triplets, etc., and the relations among them, could ultimately generate a vast number of attributes, many of which could have been diagnostic of category membership. The presence of so many alternatives would make it unlikely that category learners would select the same attributes in learning the categories. Consequently, their level of intersubject agreement did not increase, and may have decreased as a result of category learning.

Despite the general disagreement among subjects, we obtained evidence that category learning resulted in patterns belonging to the same category sharing parts that were similar in orientation. This evidence was a consequence of there being at least some agreement among subjects. It was not logically necessary for there to have been any agreement; each subject could conceivably have discovered unique attributes that were shared by members of the same category. From the point of view of our research effort, we were fortunate that there was some agreement that emerged despite the tendency toward increased intersubject variability; otherwise we wouldn't have detected the presence of perceptual learning. We were likewise fortunate to have analyzed subunit-orientation; otherwise we wouldn't have hit upon the shared attribute discovered by some of our subjects.

In conclusion, it would be reasonable to ask why investigators, beginning with Rosch and Mervis (1975), have been so successful at obtaining high levels of intersubject agreement regarding the attributes of objects, concepts, scenes, and events. It may be that levels of intersubject agreement in these studies have been overestimated because of the verbal listing technique these investigators used to identify the attributes. We argued previously that the level of intersubject agreement regarding the attributes of a stimulus depends on the number of attributes that were potentially diagnostic of category membership prior to category learning. A second, related factor concerns the extent to which the measurement technique constrains what subjects can indicate about a stimulus' attributes. While all measurement techniques, including our own, are to some extent constraining, the verbal listing technique may be excessively constraining in that it restricts the subjects to identifying attributes which lend themselves to brief verbal description. As a result, the number of alternative attributes that are verbally associated with an object, concept, scene, or event will be relatively small, and the potential for intersubject agreement is enhanced. Whether subjects agree or not may tell us more about the constraints inherent in the attribute-identification technique than it tells us about their category knowledge.

References

Banks, W. P. & Prinzmetal, W. (1976). Configurational effects in visual information processing. Perception & Psychophysics, 19, 361-367.

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. Cambridge, England: Cambridge University Press.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In Rosch, E. & Lloyd, B. C. (Eds.), Cognition and categorization. New York: Erlbaum.

Ellis, H. C. (1973). Stimulus encoding processes in human learning and memory. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 7). New York: Academic Press.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. Journal of Experimental Psychology: Learning, Memory, & Cognition, 10, 234-257.

Gibson, E. J. (1969). Principles of perceptual learning and development. New York: Appleton-Century-Crofts.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. Psychological Review, 62, 32-41.

Hock, H. S., Smith, L. B., Escoffery, L., & Bates, A. (1986). The perception of pattern: Coding the position of component elements. Submitted.

Hock, H. S., Tromley, C., & Polmann, L. (1986). Element-correspondence in the acquisition of visual categories. Submitted.

Homa, D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. Journal of Experimental Psychology: Human Learning and Memory, 5, 11-23.

Medin, D. L., Dewey, G. I. & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 607-625.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classifiction learning. Psychological Review, 85, 207-238.

Medin, D. L. & Smith, E. E. (1981). Strategies and classification learning. Journal of Experimental Psychology: Human Learning and Memory, 7, 241-253.

Murphy, G. L. & Wright, J. C. (1984). Changes in conceptual structure with expertise: Differences between real-world experts and novices. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 144-155.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. Cognitive Psychology, 9, 441-474.

Rifkin, A. (1985). Evidence for a basic level in event taxonomics. Memory & Cognition, 13, 538-556.

Rosch, E. & Mervis, C. B. Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8, 382-439.

Tversky, B. & Hemenway, K. (1983). Categories of environmental scenes. Cognitive Psychology, 15, 121-149.

Tversky, B. & Hemenway, K. (1984). Objects, parts and categories. Journal of Experimental Psychology: General, 113, 169-193.

Author Notes

Footnotes

[1]We are grateful to John Jonides for his valuable suggestions concerning the methodology used to establish the relative location of matching subunits in different patterns. The location of each subunit was determined relative to the centroid of its pattern in order to provide a frame of reference that would adjust itself to differences in the distribution of dots within each pattern.

[2]We thank Michael Lilie for writing the computer program used to analyze for common-attribute salience in subjects' parsing data.

[3]In addition to being physically larger, the subunits circled by subjects following category training tended to incorporate more dots than was the case in the no-training condition. Whereas the percentage of subunits with one or two dots remained virtually constant at 30%, the percentage of subunits with four to seven dots increased from 27% in the no-training condition to 31% and 36% in the concept-formation and paired-associate conditions, respectively. The average number of dots in each subunit was relevant to our measurement of intersubject agreement because it affects the chance rate of agreement; increases in the number of dots per subunit reduce the number of alternative ways in which a pattern can be parsed into three subunits (this is the case regardless of whether the subunits overlap, as they could in all conditions). Although this would have biased our results toward greater intersubject agreement for the training conditions, we found that intersubject agreement decreased rather than increased as a result of training.

[4]Similar analyses of two other attributes, the number of dots in each subunit and the shape of each subunit, indicated that changes in common-attribute salience as a result of category learning were no different for sets of four patterns drawn from the same category and sets of four patterns drawn from different categories.

[5]We considered making the instructions more parallel to those used in the training conditions of Experiment 1 by telling subjects that the patterns belonged to different categories. We decided not to do so because introducing the idea of category membership would have required instructional elaborations that might have distracted subjects from the main purpose of the experiment. Similar considerations influenced the choice of instructions for Experiment 3.

[6]Medin, et al. (1978), using a paired-associate procedure somewhat different from the one used in Experiment 1, have reported a case in which learning to identify individual category members was inconsistent with the abstraction of shared attributes. The use of a wide variety of training techniques and stimulus materials would be required in order to fully map the relationship between the abstraction of shared and item-specific attributes.

Table 1

Assignment of response labels to visual categories for concept-formation and paired-associate training with either the (1-4) or (5-8) subset of category members.

Concept-Formation Training

Category Members

Category	1(5)	2(6)	3(7)	4(8)
A	red	red	red	red
B	green	green	green	green
C	blue	blue	blue	blue
D	yellow	yellow	yellow	yellow

Paired-Associate Training

Category Members

Category	1(5)	2(6)	3(7)	4(8)
A	red	green	blue	yellow
B	Paris	Rome	London	Madrid
C	penny	nickel	dime	dollar
D	April	June	March	August

Table 2

Experiment 1: Percent classification and recognition errors for previously seen (original), novel, and base patterns following either concept-formation or paired-associate training.

Training Procedure	Test Patterns	Classification Errors	Recognition Errors
Concept-Formation	Bases	12.5	56.7
	Originals	5.1	10.9
	Novels	32.4	20.2
Paired-Associate	Bases	15.6	39.5
	Originals	3.7	7.4
	Novels	34.0	14.3

Figure Captions

Figure 1. The four base patterns and the eight members of each category that were generated from these base patterns.
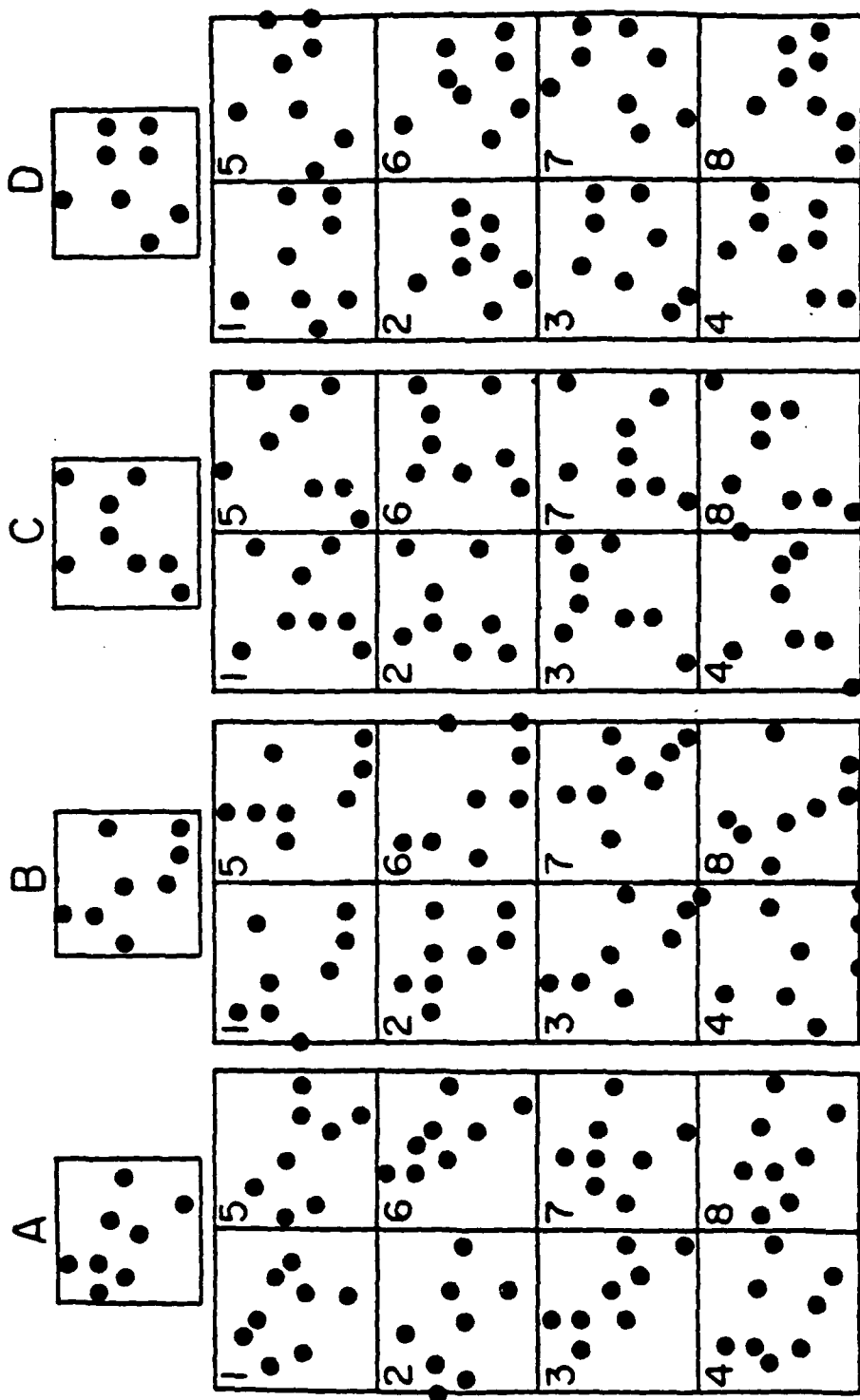
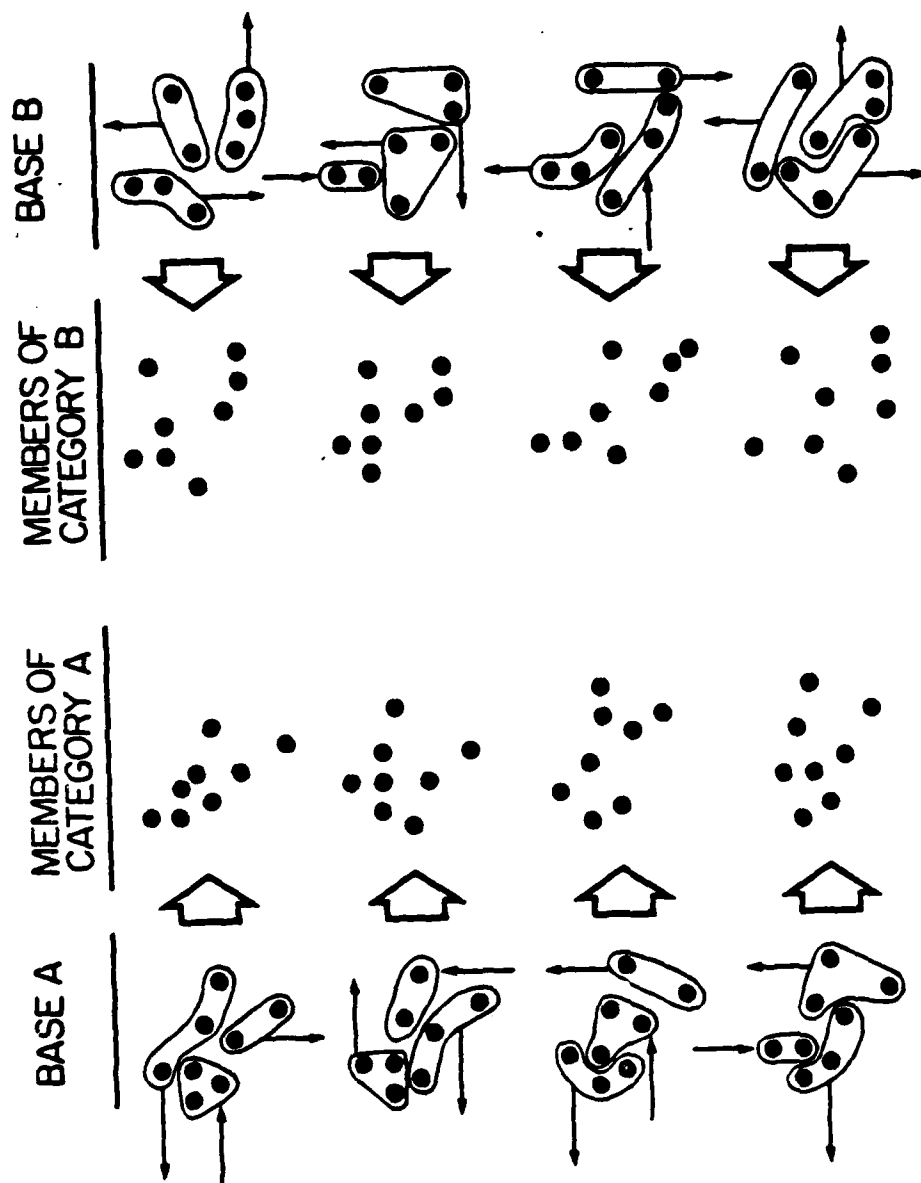Figure 2. The procedure for generating the membership of each category.

Figure 3. Experiment 1: The percentage of between-category errors, per trial block, during concept-formation and paired-associate training.
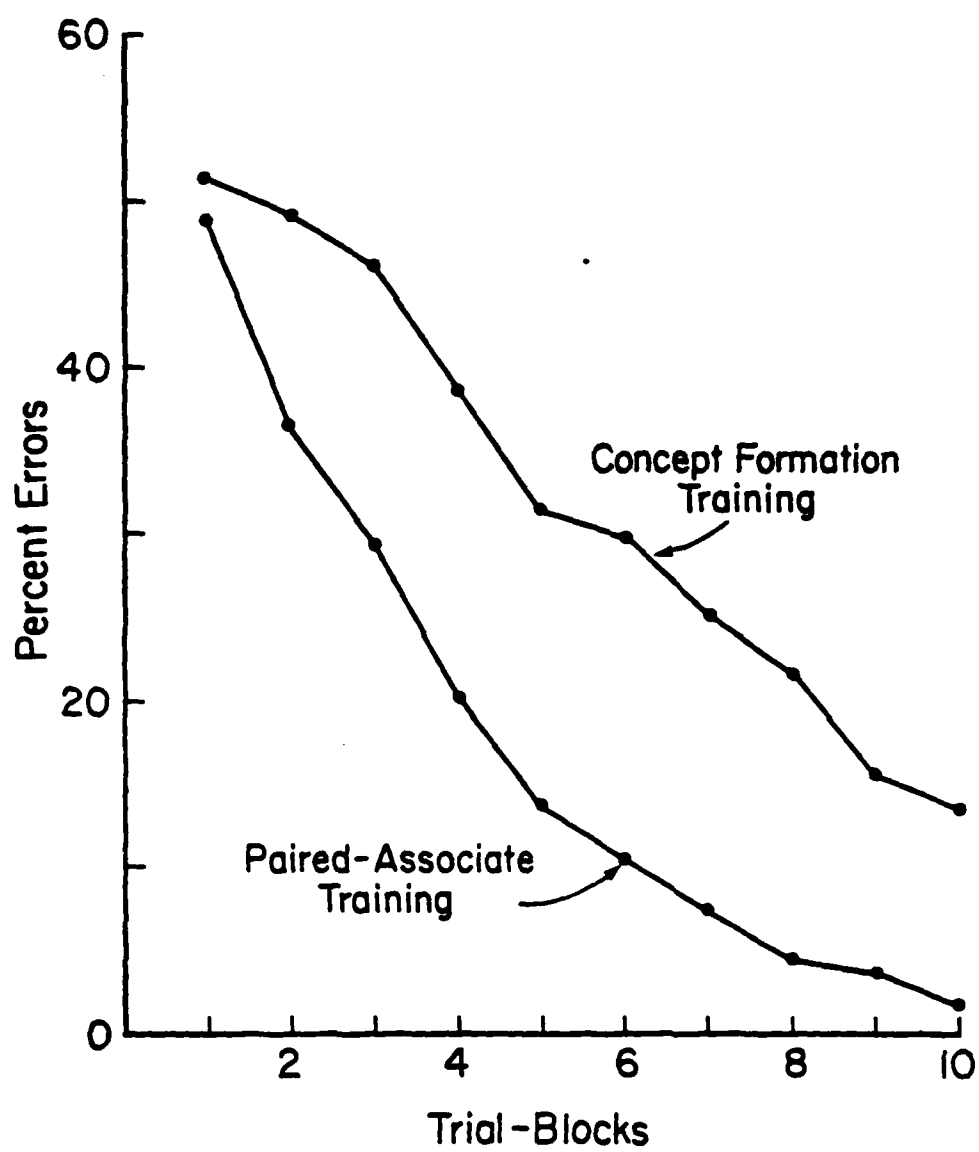
Figure 4. Experiment 1: Common-attribute salience for sets of four patterns that were members (or potential members) of the same category (within) or different categories (between). Communality was based on corresponding subunits being similar in orientation and in the same relative location for all four patterns (the latter was determined over a wide range of relative-location criteria). Subjects in the no-training condition looked for "perceptually natural" subunits in parsing each pattern. Subjects in the concept-formation and paired-associate conditions looked for subunits reflecting resemblances with other category members in parsing each pattern.
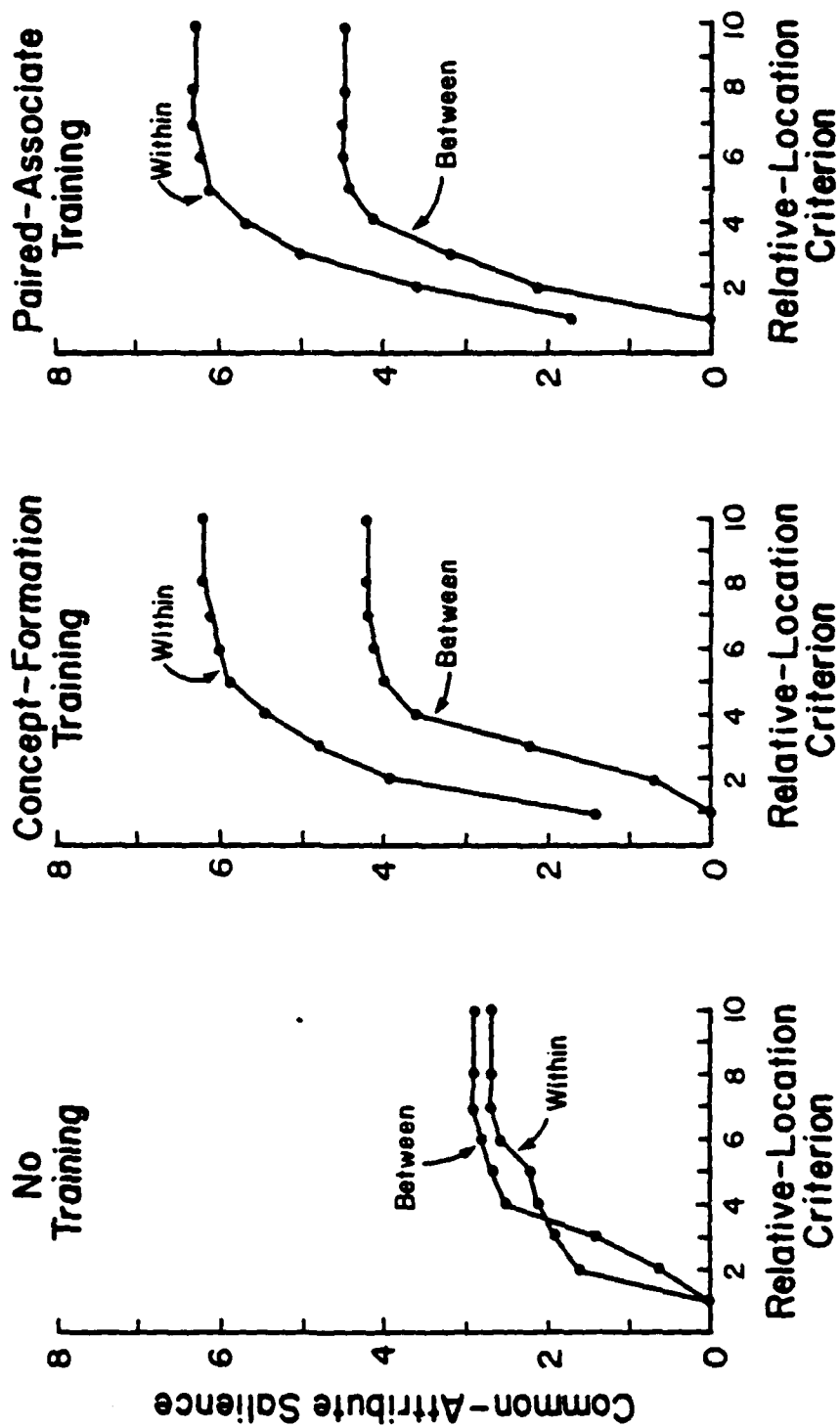
Figure 5. Experiment 2: Common-attribute salience for sets of four patterns that were potential members of the same category (within) or different categories (between). Communality was based on corresponding subunits being similar in orientation and in the same relative location for all four patterns (the latter was determined over a wide range of relative-location criteria). Subjects looked for subunits reflecting differences among the four patterns that they parsed.
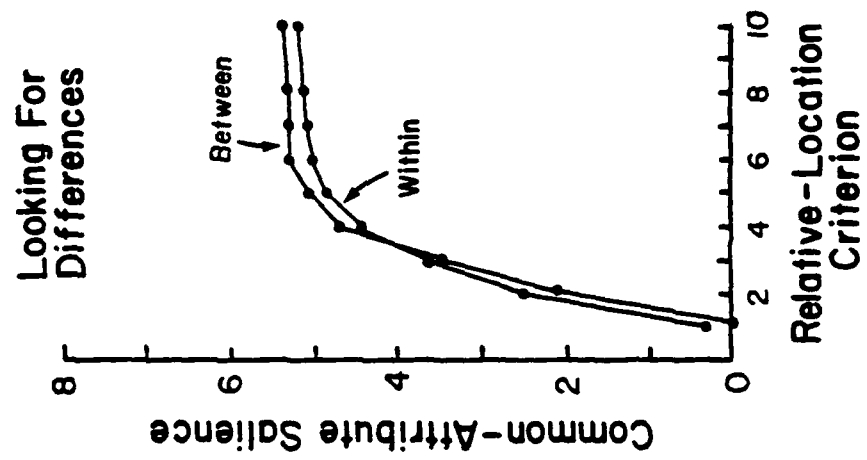
Figure 6. Experiment 3: Common-attribute salience for sets of four patterns that were potential members of the same category (within) or different categories (between). Communality was based on corresponding subunits being similar in orientation and in the same relative location for all four patterns (the latter was determined over a wide range of relative-location criteria). Subjects looked for subunits reflecting similarities among the four patterns that they parsed.
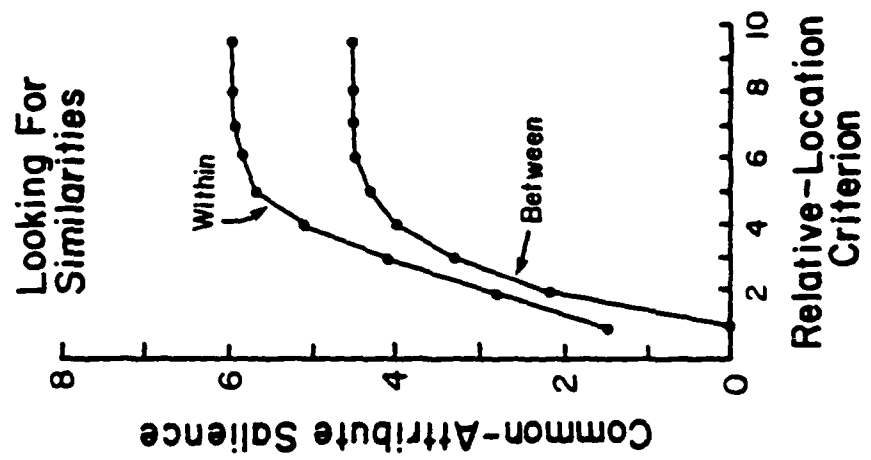
BASE B

MEMBERS OF CATEGORY B

MEMBERS OF CATEGORY A

BASE A

Looking For Differences

Looking For Similarities

Common-Attribute Salience

Within

Between

Relative-Location Criterion

The category effect in visual search:
Practice effects on catch trials

HOWARD S. HOCK, ALAN ROSENTHAL, and PHILIP STENQUIST
Florida Atlantic University, Boca Raton, Florida

Two experiments involved searching for digit or letter targets in displays of letters. On a catch trial following between-category search (for a digit among letters), a digit other than the target was presented in the display. The occurrence of incorrect target-present responses and slow, correct target-absent responses on this catch-trial decreased as the amount of preceding practice increased. This practice effect was not accompanied by shifts in decision criteria predicted from explanations of the category effect that stress between- and within-category differences in physical resemblance. Also, the effect of practice on catch trials was not accompanied by changes in speed of category-level vs. specific-level identification predicted from level-of-identification explanations of the category effect. An alternative explanation was proposed. It distinguished between attention to attributes shared by members of the target's category (resulting in the catch-trial effect) and attention to attributes specific to the target (resulting in its elimination as a function of practice).

The alphanumeric category effect refers to evidence that subjects can search through a display for an alphanumeric target at a faster rate when the nontarget items in the display belong to a different category from that of the target than when they belong to the same category as the target. This difference in search rate is indicated by a smaller effect of display size on target-present and target-absent reaction time for between-category than within-category search.

The original explanation for the category effect focused on the level of identification for the items in the display. Brand (1971), Ingling (1972), and Gleitman and Jonides (1976) proposed that between-category search rate was faster than within-category search rate because category-level identification of the target is easier than specific-level identification. As a result of this assumed difference in difficulty, Jonides and Gleitman (1972) and Egeth, Jonides, and Wall (1972) have suggested that the category-level identification of the target can be based on parallel processing of display items (as in between-category search), whereas the specific-level identification of the target requires serial processing of the display items (as in within-category search). However, evidence that specific-level identification is faster (Dick, 1971) and more accurate (Nickerson, 1973) than category-level identification directly contradicts the assumption that category-level identification is easier.

More recent explanations of the category effect have focused on how the discriminability of the target from nontarget items in the display affects the speed with which items can be rejected as nontargets. Gleitman and Jonides (1978) have argued that display items represented by both category-level and specific-level codes can be more easily rejected as nontargets when they differ from the item being searched for by both the category-level and specific-level codes (as in between-category search) than when they differ by only the specific-level code (as in within-category search). Duncan (1983) has argued that the alphanumeric category effect is due to differences in between-category and within-category resemblance. That is, on the average, members of the digit and letter categories resemble members of their own category more than they resemble members of the other category. As a result, between-category search would be faster than within-category search because items in the visual display are easier to reject as nontargets when they are easily discriminated from the target (as in the between-category condition) than when they are relatively difficult to discriminate from the target (as in the within-category condition). Duncan's (1983) physical-resemblance explanation of the category effect was based on his failure to replicate Jonides and Gleitman's (1972) oh-zero effect and Corcoran and Jackson's (1977) evidence that the difference in search rate for between-category and within-category conditions was eliminated when both conditions involved targets that were difficult to discriminate from nontarget items in the display. The latter result, which has also been reported in a recent experiment by Krueger (1984), provides strong support for Duncan's (1983) explanation of the category effect, but is not consistent with explanations that assume that performance in the visual search task is based on a search through category-level and specific-level memory

74   HOCK, ROSENTHAL, AND STENQUIST

codes for the items in a display (Francolini & Egeth, 1979; Gleitman & Jonides, 1978). Implicit in these memory-search explanations is the assumption that the performance differences in between-category and within-category search do not depend on differences in the relative difficulty of category-level or identity-level coding of display items. This implies that matching between-category and within-category conditions in target-to-nontarget resemblance should not eliminate the category effect. The results of Corcoran and Jackson's (1977) and Krueger's (1984) experiments indicate that this is not the case.

The experiments reported in this paper examined a phenomenon associated with the category effect, namely subjects' tendency to make false-alarm errors on catch trials. In standard between-category search, there are no trials in which the display includes an item that belongs to the same category as the target(s) without that item's corresponding to one of the target characters specified prior to the display. The catch trial violates this rule. For example, if the targets specified in between-category search were "2" and "4," the display might include a "3" among an array of letters. When Gleitman and Jonides (1976) presented such a catch trial (there were six items in their catch-trial display) after 192 standard between-category trials, they found that 14 of 16 subjects made false-alarm responses. In contrast, no subjects in the within-category condition made a false-alarm response when the 193rd trial included a nontarget item that had previously been used only as a target.

The catch-trial effect, like the difference in the rate of between-category and within-category search, can be explained by either the level at which display items are identified or the physical resemblance between target and nontarget items. The level-of-identification explanation would attribute the catch-trial effect in between-category search to subjects' initiating a target-present ("yes") response when they detected an item belonging to the target's category, without first identifying that item at the specific level. The physical-resemblance explanation would attribute the catch-trial effect to the perceptual confusion of the target with the nontarget item in the display that was most difficult to discriminate from the target, viz, the item that belonged to the target's category. Prior to the catch trial, most of the nontarget items in the between-category condition were dissimilar to the target (they belonged to a different category). Subjects could therefore adopt a relatively loose criterion in deciding whether items in the display corresponded to the target, resulting in a high likelihood of their making a false-alarm response on the unexpected catch trial.

The experiments reported in this study examined the effect of practice on the catch-trial effect. Regardless of whether practice increases or decreases the effect, we could examine the implications of this change for performance on the standard trials preceding the catch trial. According to the physical-resemblance explanation of the category effect, a change in the rate of false alarms on the catch trial would be due to a change in decision

criterion. If practice results in the development of a less stringent decision criterion, subjects' rates of false alarms would increase. This change should be accompanied by a decrease in the time spent processing each display item on the standard trials preceding the catch trial. The latter would be indicated if practice speeded up the search rate on standard between-category trials. Slower search rates and fewer false alarms on catch trials would be expected if a more stringent decision criterion developed with practice.

Our method of evaluating the level-of-identification explanation was somewhat different. In Experiment 2, we used the difference in reaction time between particular-digit search (e.g., look for a "4") and any-digit search (look for any digit) to infer specific-level target identification. According to the level-of-identification explanation, this difference should decrease (indicating category-level target identification) as the catch-trial effect increases.

EXPERIMENT 1

Duncan's (1983) version of the physical-resemblance explanation was proposed only for the case in which one target character is specified prior to each trial. He argued that, when more than one character is specified prior to a trial, performance in between-category search is based on the category-level identification of the display items. Since Gleitman and Jonides's (1976) catch-trial data were obtained when two targets were specified prior to each display, the starting point for our research was to test the catch-trial effect under conditions in which each display was preceded by the specification of only one target. If the catch-trial effect had not been obtained, it would have provided strong support for Duncan's (1983) argument that category-level identification does not occur when between-category search involves only one target. Since the catch-trial effect was obtained, we had the opportunity to observe whether practice would influence the size of the catch-trial effect.

Method

Stimuli. The experiment involved both between-category and within-category search. Subjects working in the between-category condition looked for a single target drawn from a set of eight possible digits: 2 through 9. Subjects working in the within-category condition looked for a single target drawn from a set of eight possible target letters: A, B, G, L, P, R, S, and Z. The nontarget items in both between-category and within-category displays were drawn from the set: C, D, E, F, H, J, K, M, N, O, V, Y.[1] The target and nontarget characters were the same as those used by Gleitman and Jonides (1976).

Each display comprised 1, 2, 4, or 6 characters, whose possible locations were defined by the 12 locations of an imaginary clock-face. For displays that included a target character, each of the eight targets in either between-category or within-category search was presented equally often at each of the 12 clockface positions, yielding 96 target-present displays. Nontarget letters were assigned to each

target-present display to create the display sizes of 1, 2, 4, or 6 characters (there were 24 of each display size in the target-present condition). The nontarget letters were randomly selected from the set of 13 nontarget letters indicated above. They were assigned randomly to various clockface locations, but with the following constraints: (1) There were no repetitions of nontarget letters within a single display, (2) one nontarget letter was always placed in the position diametrically opposed to the target (except for the one-item display) in order to maintain a constant visual angle, and (3) each nontarget letter appeared equally often. This set of 96 target-present displays was matched by another set of 96 target-absent displays that were created by substituting, for each target, a randomly selected letter from the set of 13 nontarget letters (once again with the restriction that the same character could not appear more than once in each display). This resulted in a total set of 192 displays, each involving black characters (Univers 53) presented on a white background.

The stimulus presented immediately after the experimental sequence was the catch trial, which was identical for each subject. There were six display items on this trial. In the between-category condition, the target specified prior to the display was a "3" and the digit presented in the display was a "5." In the within-category condition, the target specified prior to the catch-trial was an "A" and the display included the letter "B." The latter represented the first instance in which a letter that had theretofore been used only as a target was being used as a nontarget item in a display.

Design. Subjects were required to respond "yes" when the target specified prior to each trial appeared in the display. Otherwise, they were to respond "no." With the exception of the catch trial at the end of the experimental session, if the specified target was not present in the display, no other member of the target set was present.

The set of 96 target-present and 96 target-absent stimuli was ordered randomly within four blocks of 48 trials. Matching sequences were generated for the between-category and within-category conditions. Represented within each block of 48 were an equal number of target-present and target-absent displays for each of the four display sizes, which were also equally represented. Each block of 48 included, in random order, 24 displays that required "yes" responses (target-present displays) and 24 that required "no" responses (target-absent displays).

One group of 24 subjects received two blocks of 48 trials (Group 96); a second group of subjects received four blocks of 48 trials (Group 192). Within each group, 12 subjects participated in between-category search and 12 participated in within-category search. For Group 192, four orders of the four blocks of 48 trials were formed (Latin square), with each subject assigned to one of the four orders. The catch trial was presented on the 193rd trial. The two blocks of 48 trials presented to subjects in Group 96 were balanced so that all 192 displays presented to Group 192 were equally represented in the data collected for Group 96. The catch trial for Group 96 was presented on the 97th trial. As in Group 192, each of the four blocks of 48 trials preceded the catch trial equally often. The main experimental trials were preceded by 48 warm-up trials for both Group 96 and Group 192.

Procedure. At the start of the experimental session, subjects were shown the target and nontarget characters that were to be presented in the experiment. They were informed that the display size would vary in a random manner, and that a target would be verbally specified prior to each display. When the specified target was present in the display, the subjects were instructed to respond by pressing a button labeled "Yes." When the specified target was not present in the display, they were to respond by pressing a button labeled "No." The assignment of the subjects' right and left hands to the two responses buttons was balanced according to hand dominance: half the subjects pressed the "Yes" button with the dominant hand, and half pressed it with the nondominant hand.

The stimuli were back-projected onto a translucent screen by a Kodak Ektagraphic slide projector. A Uniblitz electronic shutter

limited the presentation of each slide to 200 msec. When viewed from a distance of 90 cm, each alphanumeric character intercepted a visual angle of 0.2° (horizontally). The imaginary clockface used to construct the displays was centered at the point of fixation, which was marked on the screen. If the clockface had been real, it would have intercepted a visual angle of 3.4°.

Each trial began with the experimenter verbally specifying a target. This was followed by the advance of the slide tray, which provided subjects with an auditory signal to look at the fixation point. The subjects were instructed to respond as quickly as possible, and were informed that some errors were expected. Corrective feedback was provided after each error.

Results

Mean reaction times for correct responses and percentage errors, excluding the data for the catch trials, are presented in Table 1. The reaction time results were consistent with the general alphanumeric category effect. That is, the effect of display size on reaction time was less for between-category than for within-category search. Analyses of variance on mean "yes" and "no" reaction times indicated that the interaction between display size (1, 2, 4, and 6) and search condition (between vs. within) significantly affected performance for Groups 96 and 192 [F(3,66) = 9.23, p < .001, MSe = 2,794.0, and F(3,66) = 18.71, p < .001, MSe = 2,053.7, respectively]. Also significant was the interaction between display size and response (yes vs. no) [F(3,66) = 4.90, p < .005, MSe = 1,861.2, and F(3,66) = 9.62, p < .001, MSe = 2,375.7, respectively]. The latter reflected the typical finding that the effect of display size is smaller when the target is present than when the target is absent. The analysis of variance for Group 192 also included practice as a factor; we contrasted performance for the first and second block of 96 trials. Practice did not moderate the category effect; the three-way interaction between practice, search condition, and display size was not significant [F(3,66) = 1.21, p > .05, MSe = 1,562.1]. Practice, however, did improve search rates: the interaction between practice and display size was significant [F(3,66) = 4.01, p < .05, MSe = 1,562.1]. Also, practice tended to reduce the difference in reaction time between "yes" and "no" responses: the interaction between practice and response was significant [F(1,22) = 5.43, p < .05, MSe = 2,392.3]. An examination of the error data provided no evidence that the above results were due to subjects' adopting differential speed-accuracy criteria in the various experimental conditions.

Having obtained evidence for the typical alphanumeric category effect, the main purpose of the experiment was to assess performance on the catch trial. None of the Group 96 or Group 192 subjects in the within-category condition made errors on the catch trial. For the between-category condition, however, 5 of 12 subjects in Group 96, as compared with 0 of 12 subjects in Group 192, made errors on the catch trial. (In this and other attempts, we were never able to approach Gleitman and Jonides's, 1976, rate of false alarms, which was 87.5%.) The difference in catch-trial false-alarm rates for the between-category conditions of Groups 96 and 192

76   HOCK, ROSENTHAL, AND STENQUIST

Table 1
Experiment 1: Mean Reaction Times (Milliseconds), Slope of the Mean Reaction Time Function
(Milliseconds per Item), and Percentage Errors for the Between-Category and
Within-Category Search Conditions for Subjects Receiving 96 Trials
(Group 96) and Subjects Receiving 192 Trials (Group 192)

Type of Search	Match With Target	Reaction Time						Errors			
		Display Size				Mean	Slope	Display Size			
		1	2	4	6			1	2	4	6
					Group 96						
Between	Yes	591	593	603	624	603	6.6	3.5	3.5	0.7	0.7
Category	No	623	607	667	706	651	19.1	0.7	0.7	0.7	0.7
Within	Yes	596	571	653	715	634	27.2	1.4	3.5	1.4	2.1
Category	No	646	655	712	839	713	38.6	7.6	2.8	0.7	2.1
				Group 192: First 96 Trials							
Between	Yes	590	577	590	621	595	6.9	1.4	2.1	2.4	0.7
Category	No	651	648	675	772	687	24.0	1.0	1.0	1.0	0.3
Within	Yes	607	618	683	771	670	33.5	0	0.3	1.4	1.4
Category	No	691	697	766	888	761	40.0	1.7	1.4	0.3	0.7
				Group 192: Second 96 Trials							
Between	Yes	578	565	599	611	588	8.3	2.4	3.5	4.2	1.4
Category	No	635	625	651	710	655	15.6	2.1	1.0	1.0	0.3
Within	Yes	577	598	656	679	628	21.3	0	2.4	3.5	3.8
Category	No	650	651	672	815	697	31.8	2.1	1.4	0.7	1.0

was significant by a Fisher-Yates test of exact probability (p < .02).

Performance on the catch trial was then contrasted with performance on the immediately preceding, six-item, target-absent trial (see Table 2). In the within-category conditions, subjects in Groups 96 and 192 were slightly more likely to make errors and were slightly faster on the immediately preceding standard trial than on the catch trial. These differences were probably due to small shifts in speed-accuracy criteria. Subjects in the between-category conditions did not make any errors on the standard trial preceding the catch trial. Of primary interest were their reaction times for correct "no" responses on the catch trial and the immediately preceding standard trial. For Groups 96 and 192, these responses were substantially slower on the catch trial than on the preceding standard trial. The effect of trial type (standard vs. catch trial) was significant [F(1,17) = 7.92, p < .02, MSe =

288,720], but neither the amount of preceding practice [Group 96 vs. Group 192; F(1,17) < 1.0, MSe = 283,423] nor the interaction between trial type and practice [F(1,17) < 1.0, MSe = 288,720] significantly affected the time required for correct "no" responses.[2]

Discussion

The results indicated that practice reduced the size of the catch-trial effect, but did not eliminate it. That is, the false alarms obtained when the catch trial was preceded by 96 standard trials were eliminated when the catch trial was preceded by 192 standard trials, but reaction times for correct "no" responses remained quite slow. It will be recalled from the introduction that the physical-resemblance explanation of the category effect would associate the reduction in the catch-trial effect with slower search rates (as indicated by a steeper slope of the function relating response time to display size). Both would

Table 2
Summary of Performance on Catch Trial and Immediately Preceding Standard Trial for
Experiments 1 and 2, Including the Proportions of Subjects Correctly Responding "No"
and the Mean Reaction Time (Milliseconds) of their Correct "No" Responses

	Proportion of Subjects Correctly Responding "No"		Reaction Time for Correct "No" Responses	
	Catch Trial	Preceding Standard Trial	Catch Trial	Preceding Standard Trial
	Between-Category			
Experiment 1: after 96 trials	0.58	1.00	1068	769
Experiment 1: after 192 trials	1.00	1.00	1245	643
Experiment 2: after 384 trials	0.92	0.92	628	627
	Within-Category			
Experiment 1: after 96 trials	1.00	0.75	805	725
Experiment 1: after 192 trials	1.00	0.83	725	683

78   HOCK, ROSENTHAL, AND STENQUIST

Table 3
Mean Reaction Time (Milliseconds), Slope of the Mean Reaction Time Function (Milliseconds per Item), and Percentage Errors for the Particular-Digit and Any-Digit Versions of Between-Category Search in Experiment 2

Match With Target	Target Specification	Reaction Time						Errors			
		Display Size				Mean	Slope	Display Size			
		1	2	4	6			1	2	4	6
			First 192 Trials								
Yes	Particular Digit	482	466	481	523	488	8.9	4.9	4.9	8.3	4.2
Yes	Any Digit	511	525	535	576	537	12.2	3.5	6.3	7.6	2.8
No	Particular Digit	527	552	586	605	568	15.4	3.5	6.9	3.5	4.2
No	Any Digit	539	559	629	742	617	40.7	5.6	2.1	2.8	4.9
			Second 192 Trials								
Yes	Particular Digit	443	470	453	477	461	4.6	3.5	1.4	6.3	3.5
Yes	Any Digit	505	494	519	529	512	6.1	5.6	3.5	2.8	6.3
No	Particular Digit	509	509	540	604	541	19.3	4.2	6.9	3.5	4.2
No	Any Digit	545	543	576	614	570	14.6	3.5	4.2	2.1	6.9

Note—Subjects received 384 trials (Group 384).

latter was responsible for the significant four-way interaction between search condition. display size. response condition, and practice [F(3.33) = 8.28, p < .001, MSe = 1,082.6]. An examination of errors provided no indication that the results described above were due to subjects' adopting differential speed-accuracy criteria in the various experimental conditions.

Only 1 of the 12 subjects made a false-alarm error on the catch trial presented at the end of the experimental sequence, a rate of false alarms that matched the rate obtained for the most immediately preceding six-item. target-absent display (with the target specified as a particular digit). As can also be seen in Table 2. the mean reaction times for correct "no" responses were virtually identical for the catch trial and the immediately preceding standard trial [t(9) < 1.0].[3] Extending the amount of practice preceding the catch trial from 192 trials to 384 trials therefore eliminated the reaction time aspect of the catch-trial effect. An examination of subjects' search rates in the particular-digit condition of Experiment 2 provided no evidence that practice also resulted in slower search. Search rates slowed with practice for target-absent responses, but speeded up by a like amount for target-present responses.

Discussion

The reaction time data meshed well with the results of previous studies. As in Taylor (1978), "yes" responses were slower in the any-digit condition than in the particular-digit condition, but search rates were the same in both conditions. Also as in Taylor (1978), this result was obtained at all levels of practice. The primary change produced by practice was the sharp reduction in search rates for "no" responses in the any-digit condition. This practice effect was in accord with the results of an experiment by Egeth, Jonides, and Wall (1972), which used practice blocks that were comparable in size to those used in this experiment (although they were presented on successive days rather than the same experimental session).

The results obtained on the catch trial continued the trend of Experiment 1. With sufficient preceding practice (384 trials), the catch-trial effect was completely eliminated. This change in the catch-trial effect with practice was not accompanied by the slower search rates that were predicted from the physical-resemblance explanation of the category effect. It was also not accompanied by the increased difference in reaction time between the particular-digit and any-digit conditions that was predicted from the level-of-identification explanation.

The additional 192 practice trials introduced in this experiment were all any-digit trials. If anything, such trials would be expected to encourage the maintenance of a strategy involving the category-level identification of display items and. thereby, the continuation of the catch-trial effect. The elimination of the catch-trial effect therefore suggested that general experience with the particular style of characters presented in the display was at least as important in eliminating the catch-trial effect as the processing demands of the search task.

GENERAL DISCUSSION

The primary experimental finding in this study was that practice reduced. and ultimately eliminated. the catch-trial effect. When the catch trial occurred after 96 standard between-category trials. subjects made either incorrect target-present responses or very slow. correct target-absent responses. When 192 trials preceded the catch trial. false alarms were eliminated. but subjects continued producing slow target-absent responses. Only when 384 practice trials preceded the catch trial were slow target-absent responses eliminated.

The results were inconsistent with Schneider and Shiffrin's (1977) claim that the category effect in visual search is the result of an automatic attention response to category-level information. Since practice tends to foster automaticity, it should have increased rather than decreased the catch-trial effect. The results were also in-

consistent with the level-of-identification explanation of the catch-trial effect. The latter would attribute the catch-trial effect to the initiation of responses on the basis of category-level identification. The reduction of the catch-trial effect with practice would be attributed to specific-level identification preceding response initiation. However, the results of Experiment 2 were inconsistent with this explanation. Reaction times in this experiment were faster for the particular-digit than for the any-digit condition at all levels of practice.

The physical-resemblance explanation of the catch-trial effect would attribute its decrease with practice to the development of more stringent criteria for determining whether a display item matched the target. There was, however, no evidence for the slowing of search rates with practice that would support the hypothesized criterion change with practice. Given that practice customarily results in faster search rates (Schneider & Schiffrin, 1977), failure to find evidence in support of the physical-resemblance explanation was not surprising. Although it is possible to speculate about additional effects of practice that might have masked the hypothesized change in criterion, it remains the case that an increase rather than a decrease in the catch-trial effect with practice would have been more amenable with Duncan's (1983) physical-resemblance explanation.

As indicated earlier, Duncan (1983) argued that within-category search is faster than between-category search because items in the display are more difficult to reject as nontargets when they resemble the target (as in within-category search) than when they do not resemble the target (as in between-category search). The results reported in this study can be accounted for by a modification of this explanation. The modification distinguishes between attributes of the target that are shared with other members of the target's category (resemblance information) and attributes that are specific to the target and therefore distinguish it from other members of its category. Accordingly, visual search would require testing each display item for the presence of attributes shared with other members of the target's category and/or attributes specific to the particular target.

Our data indicate that early in practice attention to resemblance information took precedence over attention to item-specific attributes.[4] The cost of testing each display item only for the presence of category-level resemblances was the relatively strong catch-trial effect obtained early in practice. The elimination of the catch-trial effect with further practice could then be explained by increased attention to item-specific attributes.

Why were relatively low levels of practice sufficient to reduce and eventually eliminate the catch-trial effect? Certainly 192-384 practice trials (plus 48 warm-up trials) were not enough to alter long-established differences in perceptual discriminability for alphanumeric characters. A more likely possibility is that practice provided subjects with the opportunity to "adjust" to the particular type font they were seeing. Although there was nothing

peculiar about the type font, some period of adjustment may have nonetheless been required. It might allow subjects to determine how the resemblance and item-specific attributes abstracted from previous experiences with alphanumeric characters were embodied in the particualr characters used in these experiments. Subjectively, this may have resulted in the integration of these attributes, resulting in the experience of a template-like search. Covertly, it may have resulted in the formation of decision pools that combined simultaneously available resemblance and item-specific information (Miller, 1981, 1982). The formation of such decision pools would allow for a less stringent decision criterion than would be necessary for each of the informational components working independently. On this basis, increased attention to item-specific information with extended practice would eliminate false alarms and slow responses on catch trials, and would also lead to faster search rates on noncatch trials.

The crucial evidence for the physical-resemblance explanation of the category effect was its elimination when both between-category and within-category search involved targets that resembled nontarget items in the display (Corcoran & Jackson, 1977; Krueger, 1984). However, an important factor that must be taken into consideration is that the members of most categories differ in the extent to which they are typical of their category. Rosch and Mervis (1975) have shown that the more typical a stimulus is of its category, the more it tends to resemble members of its own category and the less it tends to resemble members of contrasting categories. Stimuli that resemble members of contrasting categories as strongly as members of their own category are generally considered atypical of their category. From this point of view, the alphanumeric characters selected by Corcoran and Jackson (1977) and Krueger (1984) to match between-category and within-category resemblance were not typical members of the alphanumeric categories. When such atypical members serve as targets, attention to resemblance information is ineffective, since these targets are selected to resemble the nontarget items in the display. Target detection would then depend on attention to item-specific information, eliminating the advantage of between-category over within-category search. However, when the specified target is typical of its category, it resembles members of its own category more than members of contrasting categories. Only then would attention to resemblance information be effective in detecting the target, and an advantage be obtained for between-category search.

Our modification of Duncan's (1983) physical-resemblance explanation to account for the effects of practice on catch trials does not impair its ability to explain the various phenomena associated with the category effect. It does, however, provide a significant conceptual change. Duncan (1983) has asserted that the category effect is the result of uncontrolled differences in physical resemblance. Our modification reintroduces the idea, in-

80   HOCK, ROSENTHAL, AND STENQUIST

herent in Gleitman and Jonides' (1976) early work, that subjects' knowledge of the attribute structures for alphanumeric categories can influence visual search.

REFERENCES

BRAND, J. (1971). Classification without identification in visual search. *Quarterly Journal of Experimental Psychology*, 23, 178-186.

CORCORAN, D. W. J., & JACKSON, A. (1977). Basic processes and strategies in visual search. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.

DICK, A. O. (1971). Processing time for naming and categorization of letters and numbers. *Perception & Psychophysics*, 9(3B), 350-352.

DUNCAN J. (1983). Category effects in visual search: A failure to replicate the "oh-zero" phenomenon. *Perception & Psychophysics*, 34, 221-232.

EGETH, H., JONIDES, J., & WALL, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology*, 3, 674-698.

FRANCOLINI, C. M., & EGETH, H. E. (1979). Perceptual selectivity is task dependent: The pop-out effect poops out. *Perception & Psychophysics*, 25, 99-110.

GLEITMAN, H., & JONIDES, J. (1976). The cost of categorization in visual search: Incomplete processing of target and field items. *Perception & Psychophysics*, 20, 281-288.

GLEITMAN, H., & JONIDES, J. (1978). The effect of set on categorization in visual search. *Perception & Psychophysics*, 24, 361-368.

INGLING, N. (1972). Categorization: A mechanism for rapid information processing. *Journal of Experimental Psychology*, 94, 239-243.

JONIDES, J., & GLEITMAN, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, 12, 457-460.

JONIDES, J., & GLEITMAN, H. (1976). The benefit of categorization in visual search: Target location without identification. *Perception & Psychophysics*, 20, 289-298.

KRUEGER, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics*, 35, 536-542.

MILLER, J. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1161-1174.

MILLER, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.

NICKERSON, R. S. (1973). Can characters be classified directly as digits vs letters or must they be identified first? *Memory & Cognition*, 1, 477-484.

ROSCH, E., & MERVIS, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.

SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.

TAYLOR, D. A. (1978). Identification and categorization of letters and digits. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 423-439.

NOTES

1. Between-category search in this experiment, as well as in the experiment that follows, always involved searching for digits in a display of letters. The omission of the inverse conditions was based on the results of numerous experiments which had indicated that the typical category effect is obtained when subjects search for letters among digits or digits among letters (Duncan, 1983; Gleitman & Jonides, 1976, 1978; Jonides & Gleitman, 1972, 1976; Taylor, 1978).

2. The reaction time data reported in Table 2 included only those subjects with correct "no" responses on both the catch trial and the preceding standard trial. The analysis of variance was performed only for these subjects (7 subjects in Group 96, 12 in Group 192). The unequal numbers of subjects required the use of a least squares procedure. The inclusion of reaction times for all correct "no" responses would have only slightly altered the means reported in Table 2.

3. As in Experiment 1, the reaction time data reported in Table 2 included only those subjects with correct "no" responses on both the catch trial and the preceding standard trial (10 of the 12 subjects in Group 384). The inclusion of reaction times for all correct "no" responses would have only slightly altered the means reported in Table 2.

4. In specifying that attention to resemblance information precedes attention to item-specific information, our reasoning followed Miller's (1981) analysis of the processing of global and local pattern information. He argued that global and local information become available to a central processor over a similar time course, but attentional priority is given to global information. In a similar vein, we allow that resemblance and item-specific information also become available to a central processor over a similar time course, but attentional priority is given to resemblance information (early in practice).